

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2002-511617

(P2002-511617A)

(43) 公表日 平成14年4月16日 (2002. 4. 16)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード* (参考)
G 0 6 T 7/00	3 0 0	G 0 6 T 7/00	3 0 0 H 5 B 0 5 7
1/00	3 4 0	1/00	3 4 0 A 5 L 0 9 6

審査請求 未請求 予備審査請求 有 (全 63 頁)

(21) 出願番号 特願2000-543917(P2000-543917)  
(86) (22) 出願日 平成11年4月12日(1999. 4. 12)  
(85) 翻訳文提出日 平成12年10月12日(2000. 10. 12)  
(86) 国際出願番号 P C T / U S 9 9 / 0 7 9 3 5  
(87) 国際公開番号 W O 9 9 / 5 3 4 2 7  
(87) 国際公開日 平成11年10月21日(1999. 10. 21)  
(31) 優先権主張番号 6 0 / 0 8 1 , 6 1 5  
(32) 優先日 平成10年4月13日(1998. 4. 13)  
(33) 優先権主張国 米国 (U S)  
(31) 優先権主張番号 0 9 / 2 0 6 , 1 9 5  
(32) 優先日 平成10年12月4日(1998. 12. 4)  
(33) 優先権主張国 米国 (U S)

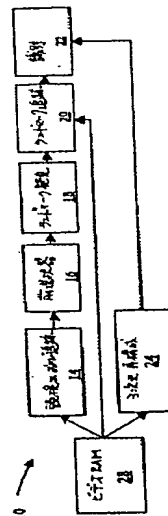
(71) 出願人 アイマティック・インターフェイス・インコーポレイテッド  
アメリカ合衆国、90404 カリフォルニア州、サンタ・モニカ、コロラド・ブールバード、2801-2803  
(72) 発明者 マウラー、トーマス  
アメリカ合衆国、90034 カリフォルニア州、ロス・アンジェルズ、ジャズミン・アベニュー、3685、ナンバー・16  
(74) 代理人 弁理士 深見 久郎 (外5名)

最終頁に続く

(54) 【発明の名称】 ビデオ画像からの顔認識

(57) 【要約】

本発明は、画像フレームにおける物体を検出および認識するための装置および関連の方法において実施される。物体は、たとえば、特定の顔の特色を有する頭である。物体検出プロセスでは、強力かつ計算上効率的な技術を用いる。物体識別および認識プロセスは、画像の特徴をジェットとして効率的に表わすモデルグラフおよびパンチグラフに基づく。ジェットはウェーブレット変換からなり、容易に識別可能な特徴に対応する画像のノードまたはランドマークローケーションで処理される。本発明のシステムは特に、人物を広範囲の姿勢角度にわたって認識する場合に好都合である。



**【特許請求の範囲】**

**【請求項1】** 画像フレームにおける物体を認識するためのプロセスであって、

画像フレームにおける物体を検出し、前記物体に関連する画像フレームの部分を囲むことと、

前記画像フレームの前記囲み部分をウェーブレット変換を用いて変換し、変換された画像を生成することと、

前記変換された画像における、代表的な複数の物体画像から生成したバンチグラフのウェーブレットジェットにより定められる物体の目立つ特徴に関連するノードの場所を求めることと、

前記物体を、物体画像のギャラリーにおける物体画像に関連するウェーブレットジェットと、前記変換された画像のノードのウェーブレットジェットとの類似度に基づいて、識別することを含む、物体を認識するためのプロセス。

**【請求項2】** 前記画像の前記囲み部分内において、前記検出された物体の大きさを定めかつその物体を中心に置いて、前記検出された物体が、前記囲み部分内で予め定められた大きさおよびロケーションを有するようにすることをさらに含む、請求項1に記載の物体を認識するためのプロセス。

**【請求項3】** 前記物体の識別前に前記物体に関連しない前記画像フレームの囲み部分の背景部分を抑制することをさらに含む、請求項1に記載の物体を認識するためのプロセス。

**【請求項4】** 前記抑制された背景部分は、前記画像フレームの囲み部分における前記物体のエッジ近くで徐々に抑制される、請求項3に記載の物体を認識するためのプロセス。

**【請求項5】** 前記物体は顔領域を示す人物の頭である、請求項1に記載の物体を認識するためのプロセス。

**【請求項6】** 前記バンチグラフは前記物体の3次元表現に基づく、請求項1に記載の物体を認識するためのプロセス。

**【請求項7】** 前記ウェーブレット変換は、ハードウェア適応位相表現を用いて行なわれる位相計算を利用して行なわれる、請求項1に記載の物体を認識す

るためのプロセス。

【請求項8】 前記場所を求めるステップは、粗から精への方策を用いて行なわれる、請求項1に記載の物体を認識するためのプロセス。

【請求項9】 前記バンチグラフは予め定められた姿勢に基づく、請求項1に記載の物体を認識するためのプロセス。

【請求項10】 前記識別するステップは、前記物体の3次元表現を用いる、請求項1に記載の物体を認識するためのプロセス。

【請求項11】 画像フレームのシーケンスにおける物体を認識するためのプロセスであって、

前記画像フレームにおける物体を検出し、前記物体に関連する各画像フレームの部分を含むことと、

各画像フレームの前記囲み部分をウェーブレット変換を用いて変換し、変換された画像を生成することと、

前記変換された画像において、代表的な複数の物体画像から生成されたバンチグラフのウェーブレットジェットにより定められる前記物体の目立つ特徴に関連するノードの場所を求めることと、

前記物体を、物体画像のギャラリーにおける物体画像に関連するウェーブレットジェットと、前記変換された画像のノードのウェーブレットジェットとの類似度に基づいて、識別することを含む、物体を認識するためのプロセス。

【請求項12】 前記物体を検出するステップは、前記物体を、前記物体に関連する軌道に基づいて画像フレーム間で追跡するステップをさらに含む、請求項11に記載の物体を認識するためのプロセス。

【請求項13】 特定の軌道に属する眺めのシーケンスから物体の最も適切な眺めを選択する、前選択プロセスをさらに含む、請求項11に記載の物体を認識するためのプロセス。

【請求項14】 前記ノードの場所を求めるステップは、画像フレーム間でノードを追跡することを含む、請求項11に記載の物体を認識するためのプロセス。

【請求項15】 追跡されたノードを、前記ノードの位置が画像フレーム間

の予め定められた位置制約を超えて逸れた場合に、再初期化することをさらに含む、請求項14に記載の物体を認識するためのプロセス。

【請求項16】 前記予め定められた位置制約は、ノードロケーション間の相対的位置に関連する幾何学的な位置制約に基づく、請求項15に記載の物体を認識するためのプロセス。

【請求項17】 前記画像フレームはステレオ画像であり、前記検出するステップは、ディスペリティヒストグラムおよびシルエット画像を生成して前記物体を検出することを含む、請求項11に記載の物体を認識するためのプロセス。

【請求項18】 前記ディスペリティヒストグラムおよびシルエット画像は、頭の移動に関連しかつ凸検出器により検出される凸領域を生成する、請求項17に記載の物体を認識するためのプロセス。

【請求項19】 前記ウェーブレット変換は、ハードウェア適応位相表現を用いて行なわれる位相計算を利用して行なわれる、請求項11に記載の物体を認識するためのプロセス。

【請求項20】 前記バンチグラフは物体の3次元表現に基づく、請求項11に記載の物体を認識するためのプロセス。

【請求項21】 前記場所を求めるステップは、粗から精への方策を用いて行なわれる、請求項11に記載の物体を認識するためのプロセス。

【請求項22】 前記バンチグラフは予め定められた姿勢に基づく、請求項11に記載の物体を認識するためのプロセス。

【請求項23】 画像フレームにおける物体を認識するための装置であって

、  
前記画像フレームにおける物体を検出し、前記物体に関連する画像フレームの部分を含むための手段と、

前記画像フレームの前記囲み部分をウェーブレット変換を用いて変換し、変換された画像を生成するための手段と、

前記変換された画像において、代表的な複数の物体画像から生成されたバンチグラフのウェーブレットジェットにより定められる前記物体の目立つ特徴に関連するノードの場所を求めるための手段と、

前記物体を、物体画像のギャラリーにおける物体画像に関連するウェーブレットジェットと、前記変換された画像のノードのウェーブレットジェットとの類似度に基づいて、識別するための手段とを含む、物体を認識するための装置。

【請求項24】 画像フレームのシーケンスにおける物体を認識するためのプロセスであって、

前記画像フレームにおいて物体を検出し、前記物体に関連する各画像フレームの部分を囲むための手段と、

各画像フレームの前記囲み部分をウェーブレット変換を用いて変換し、変換された画像を生成するための手段と、

前記変換された画像において、代表的な複数の物体画像から生成されたバンチグラフのウェーブレットジェットにより定められる前記物体の目立つ特徴に関連するノードの場所を求めるための手段と、

前記物体を、物体画像のギャラリーにおける物体画像に関連するウェーブレットジェットと、前記変換された画像のノードのウェーブレットジェットとの類似度に基づいて、識別するための手段とを含む、物体を認識するためのプロセス。

**【発明の詳細な説明】****【0001】****【関連出願の相互参照】**

本願は、米国特許法第119条(e)(1)および米国特許法施行規則第1.78条(a)(4)に基づき、1998年4月13日出願の、「VISION ARCHITECTURE TO DESCRIBE FEATURES OF PERSONS」と題された、米国仮出願第60/081,615号の優先権を主張する。

**【0002】****【発明の分野】**

本発明は、視覚に基づく物体検出および追跡に関し、より特定のには、ビデオ画像における物体、たとえば人間の顔を検出し、その物体をリアルタイムで追跡し、識別するためのシステムに関する。

**【0003】****【発明の背景】**

最近開発された物体および顔の認識技術には、弾性バンチグラフマッチングの使用が含まれる。バンチグラフ認識技術は、解析されている画像が、画像の顔部分はその画像のほとんどの部分を占めるように分割される場合に、顔を認識するのに非常に効果的である。しかしながら、弾性バンチグラフ技術は、大きなシーンにおいて、対象となる物体がそのシーンのほんのわずかな部分しか占めない場合に、物体を信頼性高く検出しないかもしれない。加えて、弾性バンチグラフ認識技術をリアルタイムで使用するには、画像を分割するプロセスが計算上効率的でなければならない。さもなければ、この認識技術の性能上の多くの利点が得られない。

**【0004】**

したがって、ビデオ画像における物体を検出し、ビデオ画像を、バンチグラフマッチングプロセスにより計算上効率的な態様でさらに処理するために準備するための、画像処理技術が、現在非常に必要とされている。本発明はこの必要を満たしている。

**【0005】**

**【発明の概要】**

本発明は、画像フレームにおける物体を検出および認識するための装置および関連の方法において実施される。物体検出プロセスは、強力かつ計算上効率的な技術を用いる。物体識別認識プロセスは、画像の特徴をジェットとして効率的に表わすモデルグラフおよびバンチグラフに基づいた画像処理技術を利用する。本発明のシステムは特に、人物を、広範囲の姿勢角度にわたって認識するのに有利である。

**【0006】**

本発明のある実施例において、物体が検出され、その物体と関連する画像フレームの一部が、囲みボックスによって境界が示される。画像フレームの囲み部分は、ウェーブレット変換を用いて変換され、変換された画像が生成される。代表的な複数の物体画像から生成されたバンチグラフのウェーブレットジェットにより定められる物体の目立つ特徴と関連するノードの場所を、変換された画像において求める。物体は、物体画像のギャラリーにおけるある物体画像と関連するウェーブレットジェットと、変換された画像のノードのウェーブレットジェットとの類似度に基づいて、識別される。

**【0007】**

加えて、画像の囲み部分内で、検出された物体の大きさを定めこの物体を中心に置くことで、検出された物体が囲み部分内で予め定められた大きさおよびロケーションを有するようにしてもよく、物体認識の前に物体に関連しない画像フレームの囲み部分の背景部分を抑制してもよい。この物体は、顔の領域を示す人物の頭であることが多い。バンチグラフは、物体の3次元表現に基づくものでもよい。さらに、ウェーブレット変換を、ハードウェア適応位相表現を用いて行なう位相計算を利用して行なうことができる。

**【0008】**

本発明のある代替実施例において、物体は画像のシーケンス内にあり、物体を検出するステップはさらに、物体を、その物体と関連する軌道に基づき、画像フレーム間で追跡することを含む。また、ノードの場所を求めるステップは、ノードを画像フレーム間で追跡し、追跡されたノードを、そのノードの位置が画像フ

フレーム間の予め定められた位置制約を超えて逸れている場合に、再初期化することを含む。加えて、画像フレームはステレオ画像でもよく、検出するステップは、頭の移動に関連する凸領域を検出することを含み得る。

#### 【0009】

本発明の他の特徴および利点は、例示として本発明の原理を示す添付の図面と関連付けられた以下の好ましい実施例の記載から、明らかになるはずである。

#### 【0010】

##### 【好ましい実施例の詳細な説明】

本発明は、画像フレームにおける物体を検出および認識するための方法および関連の装置において実施される。物体は、たとえば特定の顔の特色を有する頭である。物体検出プロセスでは、強力かつ計算上効率的な技術が用いられる。物体識別および認識プロセスでは、画像の特徴をジェットとして効率的に表わすモデルグラフおよびバンチグラフに基づく画像処理技術が用いられる。ジェットは、ウェーブレット変換からなり、容易に識別可能な特徴に対応する画像のノードまたはランドマークロケーションで処理される。本発明のシステムは特に、人物を種々の姿勢角度にわたって認識するのに有利である。

#### 【0011】

本発明の画像処理システムについて、図1—3を参照して説明する。物体認識プロセス10は、画像処理システム12から与えられるデジタル化されたビデオ画像データに対して働く。画像データは、人間の顔といった物体の種類の画像を含む。画像データは、単一のビデオ画像フレームでもよく、または、一連のシーケンシャルな単眼もしくはステレオ画像フレームでもよい。

#### 【0012】

顔の画像を弾性バンチグラフ技術を用いて処理する前に、本発明に従って、画像の中の頭を、頭検出および追跡プロセス14を用いて、大まかに配置する。画像データの性質次第で、頭検出モジュールは、例として、モーション、カラー、またはサイズ（ステレオビジョン）、トポロジーまたはパターンに基づく種々の視覚的経路のうち1つを用いる。頭検出プロセスでは、検出された頭の周りに囲みボックスを置くことによって、ランドマーク発見プロセスによって処理しなけ



ればならない画像領域を減じる。ヘッド検出および追跡プロセスから受けたデータに基づき、前選択プロセス16は、さらなる解析のために画像材料の最も適切な眺めを選択し、ヘッド検出を精密にして頭の画像を中心に置きかつその大きさを定める。選択された頭の画像は、弾性バンチグラフ技術を用いて個々の顔の特徴を検出するために、ランドマーク発見プロセス18に与えられる。顔のランドマークが顔の画像において発見されると、ランドマーク追跡プロセス20を用いてランドマークを追跡することができる。次に、ランドマークで抽出された特徴を、識別プロセス22によって、ギャラリー画像から抽出された対応の特徴と比較する。この画像認識プロセスの部分は、ランドマーク発見プロセスは比較的時間がかかりフレームレートが比較的高い一連の画像フレームに対しリアルタイムで実施できないことがよくあるため、好都合である。しかしながら、他方、ランドマーク追跡は、フレームレートよりも高速で行なうことができる。したがって、最初のランドマーク発見プロセスの間に、バッファを新たに入来する画像フレームで満たせばよい。ランドマークがの場所が求められると、ランドマーク追跡が開始され、処理システムは、バッファが空になるまでバッファされた画像を処理することによって、追いつくことができる。なお、前選択およびランドマーク追跡モジュールは、顔認識プロセスから省略してもよい。

### 【0013】

図3において、検出、ランドマーク発見および識別プロセスのための、認識プロセスのスクリーン出力が示される。左上の画像ウィンドウは、獲得した画像を、囲んでいる矩形により示される検出された頭とともに示している。頭の画像を中心に置き、大きさを再び定め、ランドマーク発見プロセスに与える。右上の画像ウィンドウは、ランドマーク発見モジュールの出力を、顔のランドマークのノードでマークされた顔の画像とともに示したものである。マークされた画像は、下側のウィンドウに示されている識別されたプロセスに与えられる。左端の画像は、識別のためにランドマーク発見プロセスから与えられた、選択された顔を表わす。右端3つの画像は、最も似ているギャラリー画像を、類似度の順に分類したものであり、最も似ている顔は左端の位置にある。各ギャラリー画像には、画像に関連するタグ（たとえばid番号および人物の名前）が添えられる。次に、

システムは、最も似ている顔に関連するタグを報告する。

【0014】

顔認識プロセスは、ステレオ画像に基づく3次元(3D)再構成プロセス24を用いて実現できる。3次元顔認識プロセスにより、視点に左右されない認識が得られる。

【0015】

本発明の顔認識プロセスを実現するための画像処理システム12を、図2に示す。この処理システムは、デジタルビデオ画像フレームのストリームを生成するビデオソース26からある人物の画像を受ける。ビデオ画像フレームは、処理のためにビデオランダムアクセスメモリ(VRAM)28に転送される。従来のCCDカメラにより生成されたデジタル化画像を発生しその画像をリアルタイムで30Hzというフレームレートでメモリに転送する、Matrox(登録商標、Dorval, Quebec, Canada、[www.matrox.com](http://www.matrox.com))から入手可能なMatrox Meteor IIは、申し分のない撮像システムである。画像フレームに対する典型的な解像度は、256画素×256画素である。画像フレームは、VRAMおよびランダムアクセスメモリ(RAM)32に結合された中央処理装置(CPU)30を備える画像プロセッサにより処理される。RAMは、本発明の顔認識プロセスを実現するためのプログラムコード34およびデータを記憶する。その代わりとして、画像処理システムは、特定用途ハードウェアにおいて実現されてもよい。

【0016】

頭検出プロセスについて、図4を参照してより詳細に説明する。顔の画像は、VRAM28に、単一画像36として、画像の単眼ビデオストリーム38として、または、画像の両眼ビデオストリーム40として、記憶することができる。

【0017】

単一画像の場合、処理時間は重要でないかもしれず、以下でより詳細に説明する弾性バンチグラフマッチングを用いて、顔を、その顔が画像の少なくとも10%をカバーしかつ直径が少なくとも50画素の場合に、検出できる。顔が画像の10%よりも小さいまたは複数の顔がある場合は、H. A. Rowley, S. Baluja and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection", P

proceedings Computer Vision and Pattern Recognition, 1998 に記載されているように、ニューラルネットワークに基づく顔検出器を用いることができる。画像が色情報を含む場合は、皮膚色検出プロセスを用いて顔検出の信頼度を高めることができる。皮膚色検出プロセスは、可能な皮膚の色を含むルックアップテーブルに基づくものでもよい。顔検出の信頼度を示すもので、バンチグラフマッチング中にまたはニューラルネットワークにより生まれる信頼度値を、皮膚色の画像領域に対し、高めることができる。

#### 【0018】

画像の動きについて、毎秒少なくとも10フレームの単眼画像ストリームを、特にこの画像ストリームが静止背景の前で動いている人物を一人のみ含むような場合、分析できる。頭追跡のある技術には、異なる画像を用いて、ある画像のどの領域が動いているのかを判断することが含まれる。

#### 【0019】

複眼画像に関して以下でより詳細に説明するように、頭が動くと、異なる画像が、動きシルエット内に凸領域を有することが多い。画像が静止背景の前で直立位置にある単一の人物を含む場合は、この動きシルエット技術によって容易に頭の動きの場所を求めてこれを追跡することができる。クラスタアルゴリズムにより、動いている領域をクラスタにグループ化する。最低しきい値サイズおよび直径を超える最高クラスタのトップが、ヘッドであるとみなされかつマークされる。

#### 【0020】

頭の動き検出のもう1つの有利な用途では、画像の動きにより影響される画素数が最低しきい値を超えるようなときにのみ使用されるグラフマッチングを用いる。このしきい値は、比較的時間のかかるグラフマッチング画像解析が、画像における変化が新たな綿密な解析を必要とするのに十分である場合のみ行なわれるように、選択される。たとえば、Turk et al., "Eigefaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, No. 1, p. 71, 1991 のように、雑音の多い動きシルエットの凸領域を求めるために他の技術を用いてもよい。D. J. Fleet, "Measurement of Image Velocity", Kluwer International Serie

s in Engineering and Computer Science, No. 169, 1992 に記載されているようなオプティカルフロー法は、どの画像領域が変化しているかを求めるための代替のかつ信頼できる手段をもたらすが、計算上はより集中的なものである。

#### 【0021】

図5を参照して、信頼性がありかつ高速の頭および顔検出を、ステレオ両眼ビデオ画像の画像ストリームを用いて行なうことができる（ブロック50）。ステレオ視野により、前景および背景物体間の区別が可能であり、かつ、これは、頭および手といったサイズがわかっている物体の物体サイズを求めることを考慮している。動きは、一連の画像における2つの画像間で、右画像チャンネルおよび左画像チャンネル双方の画像に異なるルーチンを与えることによって、検出される（ブロック52）。双方の画像チャンネルにおいて動いている画素に対しデイスパリティマップが計算される（ブロック54）。次に、凸検出器がデイスパリティヒストグラムを用いる（ブロック56）。これはデイスパリティに対する画素数を示す。あるデイスパリティ間隔に限定されたデイスパリティを有する画像領域が、デイスパリティヒストグラムの局所最大を調べることによって選択される（ブロック58）。局所最大に関連する画素は、動きシルエットと呼ばれる。この動きシルエットは2値画像である。

#### 【0022】

いくつかの動きシルエットを、ひとりの人物により生じるものとしては小さすぎるとして、廃棄してもよい（ブロック60）。所与の深さに関連する動きシルエットは、人物を他の動いている物体から区別することができる（ブロック62）。

#### 【0023】

動きシルエットの凸領域は（ブロック64）、図6に示すように凸検出器によって検出される。凸検出器は、シルエット内の凸領域を分析する。凸検出器は、近隣の画素を有する動きシルエットに属する画素68が、周囲の許容領域70内にあるのか、または、デイスパリティの幅72内にあるのか検査する。接続された許容領域は、周囲のどこかの部分に配置できる。凸検出器の出力は2値である。

## 【0024】

同様に、皮膚色シルエットを用いて頭および手を検出できる。動きシルエット、皮膚色シルエット、動きシルエットに適用される凸検出器の出力、および、皮膚色シルエットに適用される凸検出器の出力は、異なる4つの証拠マップを提供する。証拠マップは、ある画素が顔または手に属するという証拠を示す画像ドメインにわたるスカラ関数である。4つの証拠マップ各々は2値のものである。証拠マップは、所与のデイスパリティに対して線形的に重ね合わされ、局所最大について検査が行なわれる。局所最大は、頭または手を発見し得る位置の候補を示す。次に、予期される頭の直径を、証拠マップを生み出すデイスパリティマップの局所最大から推察する。上記のように、頭検出は、強い背景の動きがある場合合でさえうまく機能する。

## 【0025】

頭追跡プロセス（ブロック42）は、頭軌道検査を生じさせるのに用い得る頭位置情報を発生する。図7に示すように、新たに検出された頭の位置（ブロック78）を、既存の頭軌道と比較することができる。細線化を行ない（ブロック80）、複数の近い検出を、1つの代表的な検出と置換える（ブロック82）。新たな位置を検査し、時空間連続性を仮定して、新たに推定された位置が既に存在している軌道に属するかどうかを検査する（ブロック84）。時間 $t$ で得たフレームに対し発見されたすべての位置推定について、このアルゴリズムは、時間 $t-1$ での先行フレームに対し求めた最も近い頭位置推定を探し（ブロック86）、これを接続する（ブロック88）。十分に近い推定が発見できなければ、新たな頭が現われたと仮定し（ブロック90）、新たな軌道が始まる。個々の推定を軌道に接続するために、画像座標のみを用いる。

## 【0026】

すべての軌道には、漏れのある積分器を用いて更新される信頼度が割り当てられる。信頼度の値が予め定められたしきい値を下回る場合、その軌道は削除される（ブロック92）。ヒステリシスメカニズムを用いて、軌道の作成および削除を安定化させる。軌道を開始するためには（ブロック90）、軌道を削除するよりも高い信頼度の値に達していなければならない。

## 【0027】

前選択器16 (図2) の動作により、同じ軌道に属する一連の画像から、適切な画像を認識のために選択する。この選択は、ハードウェアの計算力が軌道の各画像を個々に分析するのに十分でない場合に、特に好都合である。しかしながら、利用できる計算力が発見されたすべての顔を分析するのに十分であれば、前選択器を用いる必要がない場合もある。

## 【0028】

前選択器16は、頭追跡プロセス14から入力を受け、ランドマーク発見プロセス18に出力を与える。入力は以下のとおりである。

## 【0029】

\* 2次元バイトアレイにより表わされる、サイズ256×256画素の単眼グレースケール画像。

## 【0030】

\* 画像のシーケンス番号を表わす整数。この数は同じシーケンスに属するすべての画像について同じである。

## 【0031】

\* 顔を囲む正方形の囲み矩形の左上および右下隅の画素座標を表わす4つの整数値。

## 【0032】

前選択器の出力は以下のとおりである。

\* 先行するシーケンスから選択された単眼グレースケール画像。

## 【0033】

\* 前選択器が入力として受ける矩形と比較して、より正確なやり方で顔の位置を表わす、正方形の囲み矩形の左上および右下隅の画素座標を表わす4つの整数値。

## 【0034】

図8に示すように、前選択器16は、頭追跡プロセス14で求めたのと同じ軌道に属する一連の顔候補を処理する(ブロック100)。ランドマーク発見に関して以下で説明するように、弾性バンチグラフマッチングを、対象となる物体(

たとえばある人物の頭)を含む画像のシーケンスに適用し(ブロック102)、さらなる処理のために最も適した画像を選択する(すなわちランドマーク発見/認識)。前選択器は、各画像を品質によって評価するためにグラフマッチングを適用する。加えて、マッチングの結果は、頭検出モジュールよりも正確な、顔の位置およびサイズに関する情報をもたらす。マッチング手順により発生した信頼度の値を、画像の適切度の尺度として用いる。前選択器は、画像を、その信頼度の値が現在のシーケンスにおいて今までに測定した最高信頼度値を超えている場合に、次のモジュールに与える(ブロック104-110)。前選択器は、検出された画像を囲みボックスによって囲み、その画像をランドマーク発見プロセス18に与える。次のプロセスは、入来する各画像に対し処理を開始するが、同じシーケンス内で信頼度の値(前選択器によって測定されたもの)のより高い画像が来た場合、処理を終了する。このことにより、CPUの作業負荷が増すが、予備結果が早く得られる。

#### 【0035】

このようにして、前選択器は、最も適切な1組の画像をさらなる処理のために取出す。その代わりとして、前選択器は、以下のようにして画像を評価することができる。

#### 【0036】

\*後続のモジュール(たとえばランドマーカ、識別子)は、最後の、したがって、前選択器が認めた最も確実な画像を、選択するために、シーケンスが終了するまで待つ。これにより、CPUの作業負荷は小さくなるが、最終結果(たとえば認識)が得られるまで時間的な遅れがあることが示唆される。

#### 【0037】

\*後続のモジュールは、前選択器が認めた各画像を取込み、それを個々に評価し、最終的な選択は次のモジュールに委ねる(たとえば認識信頼度により)。これによっても予備結果は早く得られる。この場合最終的な認識結果はあるシーケンス内で変化し、結局のところ認識率は向上する。しかしながら、この方策では、3つの評価選択肢の中でCPU時間量が最も多い。

#### 【0038】

頭部の顔のランドマークおよび特徴の場所を、図9に示す弾性グラフマッチング技術を用いて求めることができる。この弾性グラフマッチング技術では、捕捉された画像（ブロック140）を、図10を参照して以下でさらに詳細に説明するウェーブレット変換を用いて、ガボール空間に変換する（ブロック142）。変換された画像（ブロック144）は、当初の画像の各画素につき、ウェーブレット成分を表わす40の複素値によって表現される。次に、図12を参照して以下でさらに詳細に説明するモデルグラフのハードコピーを、変換された画像の上の、変化するモデルノード位置に位置決めし、最適類似度の位置を求める（ブロック146）。最適類似度の探索は、モデルグラフを画像の左上隅に位置決めし、ノードにあるジェットを抽出し、画像グラフとモデルグラフとの間の類似度を求めることによって行なうことができる。この探索は、モデルグラフを、画像の左上隅から始めて左から右へとスライドさせることによって続けられる（ブロック148）。顔のおおよその位置がわかると（ブロック150）、ノードを個別に動かして弾性グラフ歪みを導入する（ブロック152）。うまくマッチングする場所を求めるために、以下で述べる位相の影響を受けない類似度関数を用いる（ブロック154）。次に、位相の影響を受ける類似度関数を用いて、正確にジェットの場所を発見する。その理由は、位相がジェットの小さな変位の影響を大きくうけるからである。位相の影響を受けない類似度関数および位相の影響を受ける類似度関数については、図10—13を参照して以下で説明する。なお、図9では当初の画像に関するグラフが示されているが、モデルグラフの移動およびマッチングは、実際には変換された画像に対して行なわれる。

#### 【0039】

ウェーブレット変換について図10を参照して説明する。当初の画像をガボールウェーブレットを用いて処理し、畳込みの結果を得る。ガボールに基づくウェーブレットは、ガウスエンベロープにより変調された2次元の複合波フィールドからなる。

#### 【0040】

##### 【数1】



$$\psi_{\vec{k}}(\mathbf{x}) = \frac{k^2}{\sigma^2} e^{-\frac{\mathbf{x}^2 k^2}{2\sigma^2}} \left\{ e^{i\vec{k}\vec{x}} - e^{-\frac{\sigma^2}{2}} \right\} \quad (1)$$

## 【0041】

ウェーブレットは、ガウス窓により制限される、波数ベクトル $\mathbf{k} \rightarrow$ （記号 $\rightarrow$ は本来 $\mathbf{k}$ の上に示されるものである。以下の表記も同様。）の平面波であり、波長に対するその大きさは、 $\sigma$ によりパラメータ化される。括弧内の項はDC成分を除去する。波数ベクトル $\mathbf{k}$ の振幅は以下のようにして選択できる。式中 $\nu$ は所望の空間解像度に関連する。

## 【0042】

## 【数2】

$$k_{\nu} = 2^{\frac{\nu+2}{2}} \pi, \nu = 1, 2, \dots \quad (2)$$

画像位置  $\vec{x}$  を中心とするウェーブレットを用いてグレーレベル分布  $I(\vec{x})$  の画像からウェーブレット成分  $J_{\vec{k}}$  を抽出する。

$$J_{\vec{k}}(\vec{x}) = \int d\vec{x}' I(\vec{x}') \Psi_{\vec{k}}(\vec{x} - \vec{x}') \quad (3)$$

## 【0043】

典型的には、波数ベクトル $\mathbf{k} \rightarrow$ の空間を、解像度レベルが5（2分の1オクターブずつ異なる）および各解像度レベルでの配向が8の離散階層構造において（例として図13参照）サンプリングすることにより、サンプリングされた画像ポイント各々に対し40の複素値を生成する。（実成分および虚成分は平面波の正弦および余弦相を指す。） $\mathbf{k}$ -空間におけるサンプルは、インデックス $j = 1, \dots, 40$ で指定され、1つの画像ポイントを中心とするすべてのウェーブレット成分は、ジェット60と呼ばれるベクトルとみなされる。各ジェットは、 $\mathbf{x} \rightarrow$ を囲む領域の局所的な特徴を記述する。サンプリングの際の密度が十分であれば、画像を、サンプリングされた周波数によりカバーされる帯域内のジェットから再構成できる。このように、あるジェットの各成分は、画像のポイント $(x, y)$

で抽出されたガボールウェーブレットのフィルタ応答である。

【0044】

図11に示す162で示した画像グラフを用いて物体の外観を記述する（このコンテキストでは顔）。この番号で示されたグラフのノード164は、物体上のポイントを指し、ジェット160として示される。グラフのエッジ166は、ノード間の距離ベクトルで示される。ノードおよびエッジはグラフトポロジーを定める。ジオメトリが等しいグラフを比較することができる。2つのジェットの絶対成分の正規化ドット積が、ジェット類似度を定める。この値は照度およびコントラストの変化と無関係である。2つのグラフ間の類似度を計算するために、グラフ間の対応するジェットの類似度に対して和を取る。

【0045】

画像における人間の顔を発見するために具体的に指定されるモデルグラフ168を図12に示す。グラフにおいて数字が付されたノードは以下のロケーションを有する。

【0046】

- 0 右の瞳
- 1 左の瞳
- 2 鼻の上端
- 3 右眉の右隅
- 4 右眉の左隅
- 5 左眉の右隅
- 6 左眉の左隅
- 7 右の鼻孔
- 8 鼻尖
- 9 左の鼻孔
- 10 口の右隅
- 11 上唇の中心
- 12 口の左隅
- 13 下唇の中心

14 右耳の下部

15 右耳の上部

16 左耳の上部

17 左耳の下部

顔を表わすために、バンチグラフ170と呼ばれるデータ構造を用いる。これは上記のグラフと似ているが、各ノードに対しジェットを1つだけ設ける代わりに、ジェットのバンチ（束）全体72（バンチジェット）を各ノードに設ける。各ジェットは異なる顔の画像から得たものである。バンチグラフを形成するために、顔の画像のコレクション（バンチグラフギャラリー）を、頭の規定された位置のノードロケーションでマークする。この規定された位置はランドマークと呼ばれる。バンチグラフと画像とのマッチングの際、画像から抽出された各ジェットを、バンチグラフに設けられた対応のバンチにおけるすべてのジェットと比較し、最もよくマッチングしているもの1つを選択する。このマッチングプロセスは、弾性バンチグラフマッチングと呼ばれる。バンチグラフは、慎重に選択したギャラリーを用いて構成されると、局所的性質が大きく異なる多岐にわたる顔をカバーする。

#### 【0047】

画像における顔を発見するために、グラフが最もよくマッチングしている場所を発見するまで、画像フレームの上でグラフを移動させ、スケーリングする（バンチジェット内で最も適合するジェットは、ノードの現在の位置で画像から抽出されたジェットに最も類似する）。顔の特徴は顔ごとに異なるため、グラフはこのタスクに対してより一般的に作成される。たとえば、各ノードには、10から100の個々の顔から取り入れた、対応するランドマークのジェットが割当てられる。

#### 【0048】

グラフに相対的な歪みがある場合、幾何学的ひずみを説明する第2の項を導入してもよい。異なるまたは相補的でさえある2つのタスクに対して異なる2つのジェット類似度関数が用いられる。ジェット $J \rightarrow$ の成分が振幅 $a_j$ および位相 $\phi_j$ を用いた形式で書かれる場合、2つのジェット $J \rightarrow$ および $J' \rightarrow$ の類似度は、以

下のような、振幅ベクトルの正規化されたスカラ積である。

【0049】

【数3】

$$S(\bar{J}, \bar{J}') = \frac{\sum a_j a'_j}{\sqrt{\sum a_j^2 \sum a'^2_j}} \quad (4)$$

他方の類似度関数は以下の形式である。

$$S(\bar{J}, \bar{J}') = \frac{\sum a_j a'_j \cos(\phi_j - \phi'_j - \bar{d} \bar{k}_j)}{\sqrt{\sum a_j^2 \sum a'^2_j}} \quad (5)$$

【0050】

この関数は、2つのジェットが指す画像ポイント間の相対移動ベクトルを含む。グラフマッチング中に2つのジェットを比較するとき、これらの類似度をdに関して最大化し、ジェット位置の正確な判断を行なう。双方の類似度関数を用いるが、最初にグラフのマッチングを行なうときには位相の影響を受けない方（これは相対的位置とともに滑らかに変化する）が好まれることが多く、ジェットを正確に位置決めするときには位相の影響を受ける方が好まれることが多い。

【0051】

図14に示す粗から精へのランドマーク発見方策では、解像度がより低い画像においてノードおよびカーネルがより少ないグラフを用いる。粗いランドマーク発見に成功した後、特定の顔の特徴を正確に発見するために、解像度のより高い画像に対してより高い精度で局所化を行なうことができる。

【0052】

ガボール畳込みの応答は、通常絶対値および位相値として記憶される複素数である。なぜなら、ガボールジェットの比較は、値が実—虚ドメインにおいてではなくそのドメインで表現されている場合、より効率的に行なうことができるからである。典型的に、絶対値および位相値は、「浮動」値として記憶される。次に浮動に基づく算術を用いて計算を行なう。位相値は、 $-\pi$ から $\pi$ の範囲であり、 $-\pi$ は $\pi$ に等しいため、数の分布を図15に示すように円形の軸で表示できる。位相値がこの範囲を超えると常に、すなわち一定の位相値の加算または減算

によってこの範囲を超える場合、結果として得られる値をこの範囲内に再調整しなければならない。これは浮動—加算のみよりも計算上の手間がかかる。

#### 【0053】

通常用いられる整数表現および多くのプロセッサが提供する関連の算術は、2の補数である。この値は有限範囲を有するため、加算および減算の演算においてオーバフローまたはアンダフローが起り得る。2バイトの整数の正の最大数は32767である。1を加算することにより、実際—32768を表現する数になる。2の補数の整数の算術的振る舞いは、位相算術の要求に非常に近い。したがって、位相の値を、2バイトの整数で表わすことができる。位相値  $j$  を、図16に示すように整数値  $I$  にマッピングする。後で説明するマッチングおよび比較段中、 $-\pi$  から  $\pi$  の範囲にある値が要求されることは稀である。したがって、 $[-\pi, \pi]$  および  $[-32768, 32768]$  間のマッピングを、頻繁に計算する必要はない。しかしながら、位相加算および減算は頻繁に生じる。これは、プロセッサ適応間隔を用いてより高速で計算を行なう。したがって、この適応技術により、プロセッサの計算速度を大幅に向上させることができる。

#### 【0054】

顔の特徴およびランドマークの場所を示した後、顔の特徴を、図17および18に示すように、連続するフレームにわたって追跡できる。本発明の追跡技術により、特徴またはノードの追跡に失敗したかどうかを検出してそのノードに対する追跡プロセスを再初期化する追跡修正方法を用いることで、長いフレームシーケンスに対し強力な追跡を行なう。

#### 【0055】

画像シーケンスの画像  $I\_n$  における1つのノードの位置  $X\_n$  は、上記のランドマーク発見方法（ブロック180）を用いて画像  $I\_n$  でランドマークを発見する、または、追跡プロセスを用いて画像  $I\_n$  から  $I\_{(n-1)}$  へとノードを追跡することによりわかる。次に、このノードを、いくつかの技術のうち1つによって、画像  $I\_{(n+1)}$  における対応の位置  $X\_{(n+1)}$  まで追跡する（ブロック182）。以下で説明する追跡法により、有利に高速の動きに対処することができる。

## 【0056】

第1の追跡技術は線形動き予測を含む。新規の画像  $I_{n+1}$  における対応のノード位置  $X_{n+1}$  に対する探索は、動き予測子が発生した位置で開始される。速度が一定であると仮定して、先行する2つのフレーム間でのノードの変位を表わすデイスパリティベクトル  $(X_n - X_{n-1})$  を計算する。デイスパリティまたは変位ベクトル  $D_n$  を位置  $X_n$  に加算して、ノード位置  $X_{n+1}$  を予測することができる。この線形動きモデルは、一定速度の動きに対処するのに特に好都合である。線形動きモデルはまた、追跡されている物体の加速度と比較してフレームレートが高い場合に、優れた追跡をもたらす。しかしながら、フレームレートが低すぎて、画像シーケンスにおけるフレーム間で物体の加速度が大きい場合、線形動きモデルの性能は低い。動きモデルがこういった条件下で物体を追跡するのは困難であるため、フレームレートのより高いカメラを用いることが推奨される。

## 【0057】

線形動きモデルが発生する推定動きベクトル  $D_n$  が大きすぎると、動き予測における誤差の累積につながり得る。したがって、線形予測を減衰定数  $f_D$  を用いて減衰することができる。結果として得られる、予測された動きベクトルは、 $D_n = f_D * (X_n - X_{n-1})$  である。適切な減衰定数は0.9である。先行するフレーム  $I_{n-1}$  がない場合、たとえば、ランドマーク発見直後のフレームに対し、予測される動きベクトルは0と等しくなるように設定される ( $D_n = 0$ )。

## 【0058】

1次元に適用される、ガウス画像ピラミッドに基づく追跡技術は、図19に示される。当初の画像解像度を用いる代わりに、画像を2から4倍にダウンサンプリングして画像のガウスピラミッドを作成する。4レベルの画像ピラミッドは、結果として、最も粗いレベルにおいて3画素のみとして表わされている最も精細な当初の解像度レベルに対し24画素の距離をもたらす。ピラミッドのいずれかのレベルでジェットを計算および比較することができる。

## 【0059】

一般的に、ガウス画像ピラミッドにおけるノードの追跡は、最初に最も粗いレベルで行ない次に最も精細なレベルに移る。ジェットは、上記のように減衰された線形動き予測  $X_{n+1} = (X_n + D_n)$  を用い、位置  $X_{n+1}$  の実際の画像フレーム  $I_{n+1}$  の最も粗いガウスレベルで抽出され、先行する画像フレームの最も粗いガウスレベルで計算した対応するジェットと比較される。これらの2つのジェットから、デイスパリティを求める、すなわち  $X_{n+1}$  から、先行するフレームからのジェットに最もよく対応する位置までを指し示す2次元ベクトルを求める。この新しい位置が  $X_{n+1}$  に割当てられる。このデイスパリティ計算については、以下でさらに詳細に説明する。最も粗いガウスレベルの位置  $X_{n+1}$  に対応する、実際の画像の次に精細なガウスレベルの位置が  $(2 * X_{n+1})$  である、この次に精細なレベルのデイスパリティ計算についての開始点である。この点で抽出されたジェットは、先行する画像フレームの同じガウスレベルについて計算された対応するジェットと比較される。このプロセスは、最も精細な解像度レベルに達するまで、または、先行するフレームの位置に対応するノードの位置を求めるのに特定されたガウスレベルに達するまで、すべてのガウスレベルに対して繰返される。

#### 【0060】

ガウス画像ピラミッドの代表的な2つのレベルが図19に示されており、粗い方のレベル194は上であり、精細な方のレベル196は下である。各ジェットは、2つの周波数レベルに対するフィルタ応答を有すると仮定される。粗い方のガウスレベルの位置1、 $X_{n+1} = X_n + D_n$  で始めて、最低周波数ジェット係数のみを用いた第1のデイスパリティ移動により、位置2に至る。双方の周波数レベルのすべてのジェット係数を用いた第2のデイスパリティ移動により位置3に至り、これは、このガウスレベルの最終位置である。精細な方のガウスレベルの位置1は、座標が2倍の粗い方のレベルの位置3に対応する。デイスパリティ移動シーケンスは繰返され、最も精細なガウスレベルの位置3が、追跡されたランドマークの最終位置である。

#### 【0061】

実際の画像フレームにおいて追跡されたノードの新規の位置を求めた後、すべ

のガウスレベルにおけるジェットをこの位置で計算する。追跡されたノードを表わす、先行するフレームに対して計算されたジェットの記憶されたアレイを、現在のフレームに対して計算されたジェットの新規のアレイと置き換える。

#### 【0062】

ガウス画像ピラミッドを用いることには2つの主要な利点がある。第1に、粗い方のレベルにおける画素に関しノードの移動が当初の画像におけるものよりも遥かに小さく、これによって、大きな画像領域において徹底的に探索を行なう代わりに局所的な移動のみを行なうことによって追跡が可能になる。第2に、ジェット成分の計算が低周波数に対して非常に速い。なぜなら、この計算は、当初の解像度の画像の大きなカーネルウィンドウに対してではなく、ダウンサンプリングされた画像の小さなカーネルウィンドウを用いて行なわれるからである。

#### 【0063】

なお、対応レベルは動的に選択してもよい。たとえば、顔の特徴の追跡の場合、対応レベルを顔の実際の大きさに依存して選択してもよい。ガウス画像ピラミッドの大きさも、追跡プロセスを通して変更できる。すなわち、この大きさを、動きが速くなったときに大きくし動きが遅くなったときに小さくできる。典型的には、最も粗いガウスレベルの最大ノード移動は1から4画素の範囲に制限される。なお、動き予測は最も粗いレベルでのみ行なわれることも多い。

#### 【0064】

同じガウスレベルの所与の2つのジェット間の移動ベクトル（ディスパリティベクトル）の計算について次に説明する。連続する2つのフレーム間の変位を計算するために用いるある方法は、本来、D. J. Fleet および A. D. Jepson, "Computation of component image velocity from local phase information", International Journal of Computer Vision, volume 5, issue 1, pages 77-104, 1990、ならびに、W. M. Theimer および H. A. Mallot, "Phase-based binocular vergence control and depth reconstruction using active vision", CVGIP: Image Understanding, volume 60, issue 3, pages 343-358, November 1994 に基づき、ステレオ画像におけるディスパリティ推定に対して開発されたものである。複合フィルタ応答の位相の強い変化を用いて、サブピクセル精度で変位を



明確に計算する (Wiskott, L., "Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis", Verlag Harri Deutsch, Thun-Frankfurt am Main, Reihe Physik 53, PhD thesis, 1995 参照)。振幅  $a_j$  および位相  $\phi_j$  に関し  $j$  番目のガボールフィルタへの応答  $J$  を書くことにより、類似度関数を以下のように定めることができる。

【0065】

【数4】

$$S(J, J', d) = \frac{\sum_j a_j a'_j \cos(\phi_j - \phi'_j - d \cdot k_j)}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}} \quad (5)$$

【0066】

$J$  および  $J'$  を位置  $X$  および  $X' = X + d$  の2つのジェットとすると、変位  $d$  は、 $d$  に関する類似度  $S$  を最大化することによって発見でき、 $k_j$  は  $J_j$  を発生するフィルタと関連する波数ベクトルである。 $d$  の推定は、小さな変位、すなわちガボールジェットの大きな重なりのみに対して正確であるため、大きな移動ベクトルは第1の推定のみとして扱われ、このプロセスは以下の態様で繰返される。最初に、最も周波数レベルの低いフィルタ応答のみを用いて第1の推定  $d\_1$  が生まれる。次に、この推定を実行し、ジェット  $J$  をジェット  $J'$  の位置  $X'$  により近い位置  $X\_1 = X + d\_1$  で再計算する。次に、2つの最も低い周波数レベルを変位  $d\_2$  の推定のために用い、ジェット  $J$  を位置  $X\_2 = X\_1 + d\_2$  で再計算する。これは、用いた最大周波数レベルに達するまで反復され、2つの開始ジェット  $J$  および  $J'$  間の最終ディスパリティ  $d$  が、和  $d = d\_1 + d\_2 + \dots$  として与えられる。このようにして、最低周波数のカーネルの波長の半分まで変位を計算することができる (先のWiskott, 1995参照)。

【0067】

変位は浮動小数点数を用いて求めているが、ジェットを (整数) 画素位置のみで抽出してもよい (すなわち畳込みによって計算してもよい)。その結果、システマ的な丸め誤差が生まれる。このサブピクセル誤差  $\Delta d$  を補償するために、複

合ガボールフィルタ応答の位相を以下に従いシフトする必要がある。

【0068】

【数5】

$$\Delta\phi_j = \Delta d \cdot k_j \quad (6)$$

【0069】

こうすれば、ジェットは、あたかも正しいサブピクセル位置で抽出されたかのように見えるであろう。このように、ガボールジェットを、さらなる丸め誤差を考慮せずに、サブピクセル精度で追跡できる。なお、ガボールジェットは画像処理において実質的な利点をもたらす。その理由は、サブピクセル精度の問題は他のほとんどの画像処理方法において対処がより困難なためである。

【0070】

追跡誤差は、信頼度または類似度の値が予め定められたしきい値よりも小さいかどうかを判断することによって検出できる（図17のブロック184）。類似度（または信頼度）の値Sを計算して、2つの画像フレームにおける2つの画像領域がいかにかうまく対応しているかを、連続する画像フレーム間のノードの変位の計算と同時に、示すことができる。典型的には、信頼度の値が1に近いことは、対応度の高いことを示す。信頼度の値が1に近くなければ、画像において対応するポイントが発見されていないか（たとえばフレームレートが動いている物体の速度と比較して低すぎるため）、または、この画像領域がある画像フレームから次のフレームへとあまりにも大幅に変化しているため対応性をうまく定めることができないかである（たとえば瞳を追跡するノードに対し瞳が閉じられる）。信頼度の値があるしきい値よりも低いノードはスイッチオフされる。

【0071】

何らかの幾何学的制約が侵害されたときにも追跡誤差が検出され得る（ブロック186）。同時に多くのノードが追跡される場合、ノードの幾何学的構成が一致しているかどうかについて検査される。このような幾何学的制約はかなり緩やかでよい。たとえば、顔の特徴を追跡するとき鼻は目と口との間になければならないといったことである。代わりに、こういった幾何学的制約はかなり厳しいも

のでもよい。たとえば、追跡される顔の正確な形状情報を含むモデルといったものである。この中間の精度に対しては、制約は平面モデルに基づき得る。平面モデルでは、顔グラフのノードは平面にあると仮定される。正面図で始まる画像シーケンスについては、追跡されるノード位置を、アフィン変換により実際のフレームに変換される正面グラフの対応するノード位置と比較することができる。最適アフィン変換の6つのパラメータが、ノード位置における最小二乗誤差を最小にすることによって発見される。追跡されたノード位置と変換されたノード位置との間のずれがしきい値と比較される。しきい値よりもずれの大きいノードはスイッチオフされる。アフィン変換のパラメータを用いて、姿勢および相対的なスケール（開始グラフと比較したもの）を同時に求めることができる（ブロック188）。このように、この大まかな平面モデルによって、追跡誤差が予め定められたしきい値よりも大きくなることが確実になくなる。

#### 【0072】

追跡されたノードが追跡誤差のためにスイッチオフされた場合、このノードを、異なる姿勢を含むバンチグラフを有利に用いて正しい位置で再び活性化し（ブロック190）、修正された位置から追跡を続行することができる（ブロック192）。追跡されたノードがスイッチオフされた後、システムは、姿勢に固有のバンチグラフが存在する予め規定された姿勢に達するまで待ってもよい。それ以外の場合は、正面のバンチグラフのみが記憶されているならば、システムは、正面の姿勢に達するまで待って追跡誤差を修正しなければならない。記憶されたジェットのバンチを適合位置を囲む画像領域と比較してもよい（たとえば平面モデルからのもの）。これは追跡と同じ態様で作用するが、異なる点は、先行する画像フレームのジェットとの比較の代わりに、比較は例のバンチのジェットすべてについて繰返され最も似ているものが取込まれることである。顔の特徴はわかっているもので、たとえば、実際の姿勢、スケールおよび大まかな位置までもわかっているもので、グラフマッチングまたは画像における徹底的な探索および／または姿勢空間は不要であり、ノード追跡修正をリアルタイムで行なうことができる。

#### 【0073】

追跡訂正のために、多くの異なる姿勢およびスケールに対するバンチグラフは

不要である。なぜなら、画像面における回転およびスケールは、局所的画像領域またはバンチグラフのジェットの変換を図20に示すようにして行なうことにより、考慮されているからである。正面の姿勢に加え、深さにおける回転に対してのみバンチグラフを作成する必要がある。

#### 【0074】

再初期化プロセスの速度を、ひとつの画像シーケンス中は追跡される人物のアイデンティティが同じであるという事実を利用して、高めることができる。したがって、最初の学習セッションでは、人物の最初のシーケンスは、正面の顔の表情の全レパートリーを表わしている人物とともに取込むことができる。この最初のシーケンスを、異なる多くの人物に関する知識を含む大きな一般化されたバンチグラフに基づき上述の追跡および訂正方法を用いて、精度高く追跡することができる。このプロセスをオフラインで行ない、個人別にされた新たなバンチグラフを生成することができる。次に、個人別にされたバンチグラフを用いて、この人物をリアルタイムで高速レートで追跡できる。なぜならば、個人別にされたバンチグラフは一般化された大きなバンチグラフよりも遥かに小さいからである。

#### 【0075】

再初期化プロセスの速度を、部分バンチグラフ再初期化を利用して高めることもできる。部分バンチグラフは、全バンチグラフのノードのサブセットのみを含む。このサブセットの大きさは1つのノードぐらいでしかない。

#### 【0076】

姿勢推定バンチグラフは、画像面において定められる2次元バンチグラフの一族を利用する。1つの族内の異なるグラフは、頭の、異なる姿勢および／またはスケールを説明する。ランドマーク発見プロセスは、画像における頭の姿勢または大きさを求めるために、その族からの各バンチグラフを入力画像とマッチングさせようとする。こうした姿勢推定手順の一例が図21に示される。この姿勢推定の第1ステップは、通常のランドマーク発見のものと等しい。画像（ブロック198）は、グラフ類似度関数を用いるために、変換される（ブロック200および202）。次に、1つだけではなく、3つのバンチグラフの族を用いる。第1のバンチグラフは正面の姿勢の顔のみを含み（上記の正面図と等しい）、他の

2つのバンチグラフは4分の1だけ回転させた顔を含む（一方が左への回転を表わし、他方が右への回転を表わす）。先のように、グラフ各々に対する初期位置は左上隅にあり、グラフの位置は画像においてスキャンされ、ランドマーク発見後最も高い類似度を戻す位置およびグラフが選択される（ブロック204-214）。

#### 【0077】

各グラフに対する最初のマッチングの後、最終位置の類似度が比較される（ブロック216）。画像において与えられた姿勢に最もよく対応しているグラフは類似度が最も高い（ブロック218）。図21において、左に回転させたグラフは、類似度によって示されるように最もよく適合している。像における顔の解像度および回転の角度に応じて、正しいグラフおよび他の姿勢に対するグラフの類似度は、変化し、グラフを規定した2つの姿勢間のおよそ中間に顔があるときに、非常に近くなる。より多くの姿勢に対してバンチグラフを作成することにより、姿勢推定手順がより精細になり、これにより、他の方向（たとえば上または下）への頭の回転およびハンドル回転のより多くの角度が区別される。

#### 【0078】

カメラから任意の距離にある顔を確実に発見するために、同様の方策を用いて、各々がスケールの異なる異なる2または3のバンチグラフを用いてもよい。画像における顔は、最高のものを顔の画像に戻すバンチグラフと同じスケールを有すると仮定される。

#### 【0079】

上記の技術に関連する3次元（3D）ランドマーク発見技術を用いて、異なる姿勢に適合させた複数のバンチグラフを用いることもできる。しかしながら、この3次元方策は3次元空間で規定されるバンチグラフ1つのみを用いる。3次元グラフのジオメトリは、平均的な顔または頭のジオメトリを反映する。回転度数の異なる数人の人物の顔の画像からジェットを抽出することにより、3次元バンチグラフを作成する。これは、2次元方策と似ている。次に各ジェットを3つの回転角度でパラメータ化する。2次元方策のように、ノードは頭の表面の基準点に位置する。次に、3次元グラフの突起をマッチングプロセスで用いる。3次元

方策におけるある重要な一般化は、各ノードに、異なる姿勢に適合するバンチジェットのパラメータ化された族を設けることである。第2の一般化は、グラフが画像面での変換のみではなく3次元空間でユークリッド変換し得ることである。

#### 【0080】

3次元グラフマッチングプロセスは、粗から精への方策として定めることができる。ここでは初めにノードおよびカーネルの少ないグラフを用い、次に後続のステップではより密度の高いグラフを利用する。粗から精への方策は、顔のある領域における特徴ポイントの高精度の局所化が所望される場合に特に適している。このように、階層的方策を採用することによって、計算上の手間を省く。ここでは、最初に粗い解像度についてランドマーク発見を行ない、次に、適合したグラフをより高い解像度で検査して、ある領域をより精細に解析する。

#### 【0081】

さらに、計算上の作業負荷を、マルチプロセッサマシンで簡単に分割してもよい。こうすれば一旦粗い領域が発見されると、少数の子プロセスが並列に、各々が画像全体のうちそれ自身の部分に対して作業を開始する。子プロセスの最後に、これらプロセスはそれらが位置する特徴座標をマスタプロセスに送り、マスタプロセスは適切にこれらをスケーリングし組合せて当初の画像に合うようにし、こうして計算の総時間が大幅に短縮する。

#### 【0082】

テクスチャマッピングされた頭部3次元モデルを構成するために、多数の方法が開発されてきた。このセクションでは、ステレオに基づく方策について説明する。ステレオに基づくアルゴリズムを、十分に校正されたカメラの場合について説明する。このアルゴリズムは、画像の画素の面積に基づきマッチングを行ない、密度の高い3次元情報が必要な場合に適している。次にこれを用いてより高度な物体の記述を正確に定めることができる。ステレオ撮像およびマッチングに関するさらなる背景情報は、U. Dhond および J. Aggrawal, "Structure from Stereo: a Review", IEEE Transactions on Systems, Man, and Cybernetics, 19 (6), pp. 1489-1510, 1989, または、より最近では、R. Sara および R. Bajcsy, "On Occluding Contour Artifacts in Stereo Vision", Proc. Int. Conf. Com

puter Vision and Pattern Recognition, IEEE Computer Society, Puerto Rico, 1997.; M. Okutomi および T. Kanade, "Multiple-baseline Stereo", IEEE Trans. on Pattern Analysis and Machine Intelligence 15 (4), pp. 353-363, 1993; P. Belhumeur, "A Bayesian Approach to Binocular Stereopsis", Intl. J. of Computer Vision, 19 (3), pp. 237-260, 1996; Roy, S. および Cox, I., "Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem", Proc. Int. Conf. Computer Vision, Narosa Publishing House, Bombay, India, 1998; Scharstein, D. および Szeliski, R., "Stereo Matching with Non-Linear Diffusion", Proc. Int. Conf. Computer Vision and Pattern Recognition, IEEE Computer Society, San Francisco, CA, 1996; ならびに Tomasi, C. および Manduchi, R., "Stereo without Search", Proc. European Conf. Computer Vision, Cambridge, UK, 1996 に示される。

#### 【0083】

二眼式立体映像において重要なことは、対応性（マッチング）の問題であるとして知られている。すなわち、両眼ステレオからレンジ（領域）データを取戻すために、空間3次元ポイントの対応の射影を左右の画像で発見しなければならない。探索—空間次元を減じるために、エピポーラ（epipolar）制約を適用する（S. Maybank および O. Faugeras, "A Theory of Self-Calibration of a Moving Camera" Intl. J. of Computer Vision, 8 (2), pp. 123-151, 1992参照）。二眼式立体映像は、以下のようにして4ステッププロセスで定めることができる。

#### 【0084】

\*校正：カメラのパラメータを計算する。

\*調整：ステレオ対を投影し、画像における対応の特徴が同じ線にあるようにする。このラインは視線像（エピポーラライン、epipolar line）と呼ばれる。これは全く必要なわけではないが、アルゴリズムの性能を大幅に高める。なぜなら、マッチングプロセスを、1次元探索として、調整された画像の水平線に沿って行なうことができるからである。

#### 【0085】

\*マッチング：探索ウィンドウにおいてコスト関数を各位置について局所的に

計算する。相関の最大値を用いてステレオ対における対応の画素を選択する。

#### 【0086】

\*再構成：ステレオ対においてマッチングした画素座標から3次元座標を計算する。

マッチング誤差を取除くために、マッチングの直後に後処理を加えることができる。可能性として、誤差は、マッチングの曖昧さにより生じ、その原因はマッチングが局所的に行われることにあることが多い。いくつかの幾何学的制約およびフィルタリングを適用して、誤ったマッチングの数を減じる。連続する面を扱うときには（例として前位置にある顔）、内挿を用いてマッチングしていない領域を回復させることができる（大抵は非テクスチャ領域であり相関スコアは明らかなモノモード最大値を有さない）。

#### 【0087】

調整および再構成プロセスにおいて用いられる等式に至る形式は、射影幾何学と呼ばれており、O. Faugeras, "Three-Dimensional Computer Vision, A Geometric Viewpoint", MIT Press, Cambridge, Massachusetts, 1993 において詳細に示されている。使用されるモデルは多大な利点をもたらす。一般に、図22に示す簡単なピンホールカメラモデルを想定する。必要であれば、校正時にレンズの歪を計算することもできる（最も重要なファクタは放射方向のレンズ歪である）。実際的な観点からすれば、校正は、校正補助具すなわち既知の3次元構造の物体を用いて行われる。通常、目に見えるドットまたは正方形パターンを備える立方体を、図23に示すように校正補助具として用いる。

#### 【0088】

調整アルゴリズムを簡単にするために、最初に各ステレオ対の入力画像を調整し（N. Ayache および C. Hansen, "Rectification of Images for Binocular and Trinocular Stereovision", Proc. of 9th International Conference on Pattern Recognition, 1, pp. 11-16, Italy, 1988参照）、対応するポイントが同じ画像の線上にあるようにする。こうすれば、明らかに、対応するポイントは左右の調整された画像において座標  $(u_l, v_l)$  および  $(u_l-d, v_l)$  を有し、「d」はデイスパリティとして知られているものである。調整プロセスの詳細に



については先のFaugerasを参照。調整面（画像を投影して調整された画像を得るのに用いる面）の選択は重要である。通常、この面は、投影された画像の歪が最小になるように、かつ、対応する画素が図24に示すように同じライン番号（視線像は並列でありかつ整列している）に沿う位置にあるように、選択される。このような構造は標準ジオメトリと呼ばれる。

#### 【0089】

図26を参照して、マッチングは、左右の画像において対応するポイントを見出すプロセスである。いくつかの相関関数を用いてこのデイスパリティを計算することができる。たとえば、正規化された交差相関（H. Moravec, "Robot Rover Visual Navigation", Computer Science: Artificial Intelligence, pp. 13-15, 105-108, UMI Research Press 1980/1981参照）は、以下によって与えられる。

#### 【0090】

$$c(I_L, I_R) = 2 \text{cov}(I_L, I_R) / (\text{var}(I_L) + \text{var}(I_R)) \quad \dots (6)$$

式中、 $I_L$ および $I_R$ は調整された左右の画像である。相関関数を矩形領域のポイント $(u_L, v_L)$ および $(u_R, v_R)$ に適用する。サイズ $1 \times N$ の探索ウィンドウについて、コスト関数 $c(I_L, I_R)$ を、図25に示すようにして計算する（調整プロセスのため）。ここで $N$ はある許容可能な整数である。左の画像の各画素 $(u_L, v_L)$ について、マッチングにより、相関プロファイル $c(u_L, v_L, d)$ が生まれる。ここで「 $d$ 」はポイント $(u_L, v_L)$ でのデイスパリティとして定められる。すなわち以下のとおりである。

#### 【0091】

$$d_u = u_R - u \quad \dots (7)$$

$$d_v = 0 \quad \dots (8)$$

2番目の等式は、視線像が整列していることを表わしている。マッチング手順により、デイスパリティマップが出力される、または、ベース画像に重ねることができるデイスパリティの画像が出力される（ここではステレオ対の左の画像）。このデイスパリティマップは、「ステレオ対の右の画像において画素の対応物を

発見するために視線像に沿ってどれだけ移動すべきか」を示している。

#### 【0092】

マッチング時にいくつかの精細化を利用してもよい。たとえば、可能な対応物のリストを各ポイントで保持し、鮮明度の制約、順序の制約およびディスパリティ傾斜制約といった制約を用いて (A. Yuille および T. Poggio, "A Generalized Ordering Constraint for Stereo Correspondence", MIT, Artificial Intelligence Laboratory Memo, No. 777, 1984; 上記 Dhond et al. ならびに Faugeras 参照)、不可能な構造を取除くことができる (上記 R. Sara et al, 1997参照)。クロスマッチングを用いることもできる。このマッチングは、左から右へと次に右から左へと行われ、候補 (相関ピーク) は、双方のマッチングが同じ画像の画素に至る場合に受入れられる。すなわち、以下の場合である。

#### 【0093】

$$d_{LR} = u_L - u = -d_R \quad \dots (9)$$

式中  $d_{LR}$  は、左から右へのマッチングで発見されたディスパリティであり、 $d_{RL}$  は右から左の場合のものである。さらに、ピラミッド形の機構を用いて探索ウィンドウを制限することによりマッチングプロセス全体を助けることができる。これは、先行するレベルの推定を用いて、解像度のピラミッドの各レベルでマッチングを行なうことにより実現される。なお、階層的機構は表面の連続性を強化する。

#### 【0094】

なお、ステレオを2次元分割目的で用いた場合、ディスパリティマップのみが必要である。このようにして、先に説明した校正プロセスなしで済まし、投影ジオメトリの結果を用いることができる (Q. T. Luong, "Fundamental Matrix and autocalibration in Computer Vision", Ph.D. Thesis, University of Paris Sud, Orsay, France, December 1992参照)。これは、基本マトリックス (Fundamental Matrix) が利用できるならば調整が可能であることを示している。今度は、基本マトリックスを用いて画像を調整し、マッチングが先に述べたように実施できるようにする。

#### 【0095】

3次元位置推定を正確にするために、整数ディスプレイマップのサブピクセル補正を計算し、結果としてサブピクセルディスプレイマップを得る。サブピクセルディスプレイは、以下のいずれかにより得ることができる。

【0096】

\*検出された最大値の周囲の相関スコアの2次内挿を用いる。

\*F. Devernay, "Computing Differential Properties of {3-D} Shapes from Stereoscopic Images without {3-D} Models", INRIA, RR-2304, Sophia Antipolis, 1994に記載されているような、より一般的な方策を用いる。(ここでは、表面の平坦なパッチが撮像されると仮定して、斜めからの投影により生まれる左右の相関ウィンドウ間の歪を考慮している。)

第1の方策は最も速いが、第2の方策ではサブピクセルディスプレイのより信頼性の高い推定が得られる。高速サブピクセル推定を行なう一方で推定の正確度を保つために、以下のようにする。 $I_L$ および $I_R$ が左および右の調整された画像であるとする。 $\epsilon$ が未知のサブピクセル補正であるとし、 $A(u, v)$ が相関ウィンドウを左から右の画像へとマッピングする変換であるとする(平坦なターゲットに対しては画像列を保存するアフィンマッピングである)。左右の画像の対応する画素については以下のとおりである。

【0097】

$$I_R(u_L - d + \epsilon, v) = \alpha I_L(A(u_L, v_L)) \quad \dots (10)$$

式中、係数 $\alpha$ はカメラ利得において起こり得る差を考慮している。「 $\epsilon$ 」および「 $A$ 」に関する上記の式の1次線形近似により、線形系がもたらされ、各係数は、対応する左右の相関ウィンドウにわたって推定される。この線形系の最小二乗解により、サブピクセル補正が得られる。

【0098】

なお、連続する面を回復させる場合(前姿勢の顔について)、内挿方法を、フィルタリングされたディスプレイマップに用いることができる。このような方法は以下のことを考慮することによって導き出すことができる。ここでは下にある面が連続していると想定しているため、内挿され平滑化されるディスプレイマップ $d$ は、以下の等式を確かめなければならない。

## 【0099】

$$\min \iint [(d' - d) + \lambda (\nabla d)^2] du dv \quad \dots (11)$$

式中、 $\lambda$ は平滑化パラメータであり、積分を画像全体に対して取入れている（画素座標 $u$ および $v$ に対して）。反復アルゴリズムは、オイラー等式を用い、かつラプラシアン演算子の近似を用いて、簡単に得られる。

## 【0100】

ディスパリティマップおよびカメラ校正から、3次元ポイントの空間位置が、三角測量に基づいて計算される（上記のDhond et al.参照）。（1つの画像ステレオ対からの）再構成の結果、空間ポイントのリストを得る。

## 【0101】

いくつかの画像を用いる場合（複眼 (polynocular)ステレオ）、検証ステップを用いることができる（R. Sara, "Reconstruction of 3-D Geometry and Topology from Polynocular Stereo", <http://com.felk.cvut.cz/~sara>参照）。この手続中、すべてのステレオ対からの再構成されたポイントの組は、すべてのカメラ対のディスパリティ空間に再び戻すように投影され、投影されたポイントが各対の他の画像において予測された位置とマッチングしている場合は、検証される。検証によってはずれ値が非常に効果的に取除かれるように思われる（特に閉塞近くのマッチングの結果）。

## 【0102】

図26は、ステレオアルゴリズムを、テクスチャ光を投影することにより得られた画像のステレオ対に適用した、典型的な結果を示す。図26の上の列は、主体が動かないことを保証して短い時間間隔において取込まれた左右のおよび色画像を示している。下の列は、ステレオをテクスチャ画像に適用し、かつ色画像でテクスチャマッピングすることにより得た、再構成された顔のモデルの2つの図を示している。なお、内挿およびフィルタリングをディスパリティマップに適用しているので、顔の再構成は滑らかで連続している。また、この結果は、ステレオから得た未処理のポイントの組として表示されており、これらのポイントを網目状にして、たとえば、アルゴリズム用いて、位置を記憶されたギャラリー画像から抽出したジェットと比較し、連続面を得ることができる。顔認識アプリケー

ションのように完全なグラフを比較するか、または、部分的なグラフもしくは個々のノードを比較する。

#### 【0103】

ジェットを実際の比較のために抽出する前に、多数の画像正規化を適用する。このような正規化の1つは背景抑制と呼ばれる。プローブ画像に対する背景の影響を抑制しなければならない。なぜなら、プローブおよびギャラリー画像間で異なる背景が、類似度を低下させ、しばしば分類の誤りに至ることがあるためである。したがって、顔を囲むノードおよびエッジを顔の境界とする。背景画素は、顔から逸れているときに、滑らかにトーンダウンされる。頭の外側の各画素値は以下のように変形される。

#### 【0104】

$$P_{\text{new}} = P_{\text{old}} \cdot \lambda + c \cdot (1 - \lambda) \quad \dots (12)$$

#### 【0105】

#### 【数6】

$$\lambda = \exp\left(-\frac{d}{d_0}\right) \quad (13)$$

#### 【0106】

ここでcは、グラフの最も近いエッジからの画素位置のユークリッド距離を表わす、一定の背景グレー値である。d<sub>0</sub>は、一定のトーンダウン値である。もちろん、画素値およびグラフ境界からの距離間の、他の関数従属性も可能である。

#### 【0107】

図28に示すように、自動背景抑制により、グレー値は、最も近いエッジから逸れているときに滑らかに一定になる。この方法ではなおも顔を囲む背景領域が見えているが、画像において強く妨害するようなエッジはない。これは、この領域が単純に一定のグレー値で満たされているようなときに生じるであろう。

#### 【0108】

上記は本発明の特定の実施例に関するものであるが、当業者であれば、これが例示としてのものにすぎず、こうした実施例における変更が、本発明の原理から

逸脱することなく可能であり、本発明の範囲は前掲の特許請求の範囲によって定められていることを理解するであろう。

【図面の簡単な説明】

- 【図1】 本発明に従う顔認識プロセスのブロック図である。
- 【図2】 本発明に従う顔認識システムのブロック図である。
- 【図3】 図1の認識プロセスの検出、発見、および識別プロセスを示す、一連の画像である。
- 【図4】 本発明に従う頭検出および追跡プロセスのブロック図である。
- 【図5】 本発明に従うデイスパリティ検出プロセスを示すためのフロー図に、画像を添えたものである。
- 【図6】 本発明に従う凸検出器の概略図である。
- 【図7】 本発明に従う頭追跡プロセスのフロー図である。
- 【図8】 本発明に従う前選択器のフロー図である。
- 【図9】 図1の顔認識装置およびシステムのランドマーク発見技術を示すフロー図に写真を添えたものである。
- 【図10】 本発明に従うガボールウェーブレットを用いた顔の像の処理を示す、一連の画像である。
- 【図11】 本発明に従う、図10のウェーブレット処理技術を用いた、ジェット、画像グラフ、およびパンチグラフの構成を示す、一連のグラフである。
- 【図12】 本発明に従う、顔の画像を処理するためのモデルグラフの図である。
- 【図13】 ウェーブレット処理を用いて顔の特徴をの場所を求めることを示す2つの図を含む。
- 【図14】 粗から精へのランドマーク発見技術を示すための顔の図であり、これとともに抽出された目および口領域が示されている。
- 【図15】 相の円形の振る舞いを示す概略図である。
- 【図16】 本発明に従う、円形の振る舞いを有する相の2の補数を示す概略図である。
- 【図17】 本発明のランドマーク発見技術により発見されるランドマーク

を追跡するための追跡技術を示す、フロー図である。

【図18】 本発明に従う顔の特徴の追跡を示す、一連の顔の画像である。

【図19】 1次元でのランドマーク追跡を示すためのガウス画像ピラミッド技術の図である。

【図20】 2つの顔の一連の画像に、50の画像フレームのシーケンスにわたる顔の特徴の追跡を示す、姿勢角度対フレーム数のグラフを添えたものである。

【図21】 図1の認識装置およびシステムの姿勢推定技術を示すためのフロー図に、写真を添えたものである。

【図22】 3次元(3D)ビューアクセスの配向を示すピンホールカメラモデルのグラフである。

【図23】 3次元カメラ校正構造の斜視図である。

【図24】 同じライン番号に添うステレオ画像の対応画素を投影するための調整の概略図である。

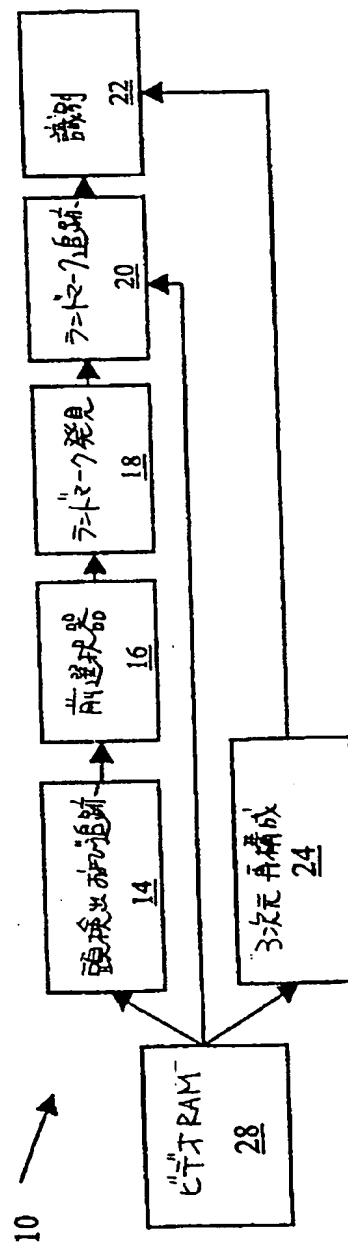
【図25】 一方の画像フレームのウィンドウと他方の画像フレームの探索ウィンドウとの間の相関マッチングプロセスを示す、画像フレームである。

【図26】 ステレオ画像対の画像、ディスパリティマップ、および3次元画像デコードを示す画像再構成である。

【図27】 本発明に従う画像識別プロセスのフロー図である。

【図28】 背景抑制の使用を示した画像である。

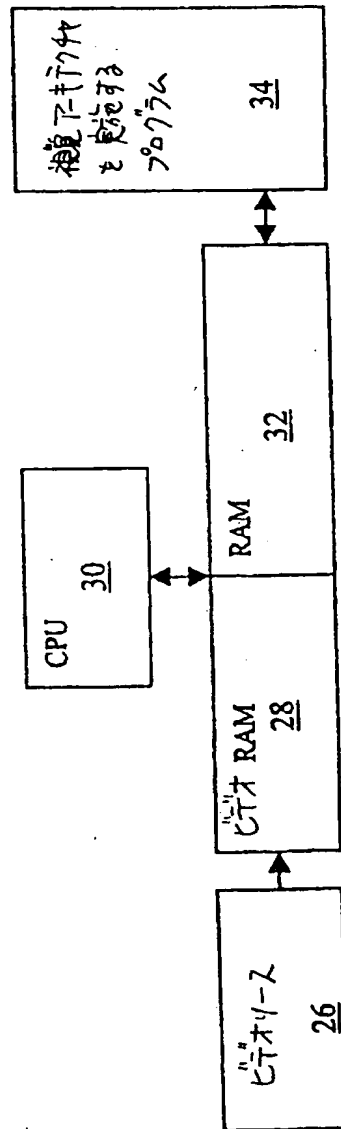
【図1】





【図2】

12



【図3】

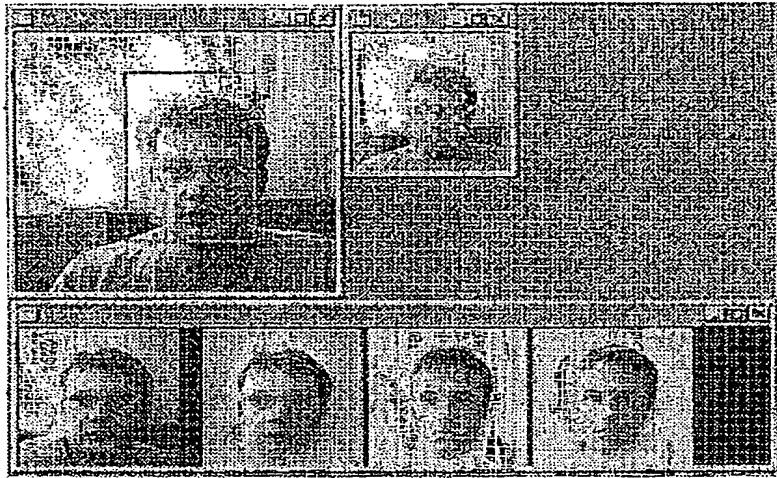
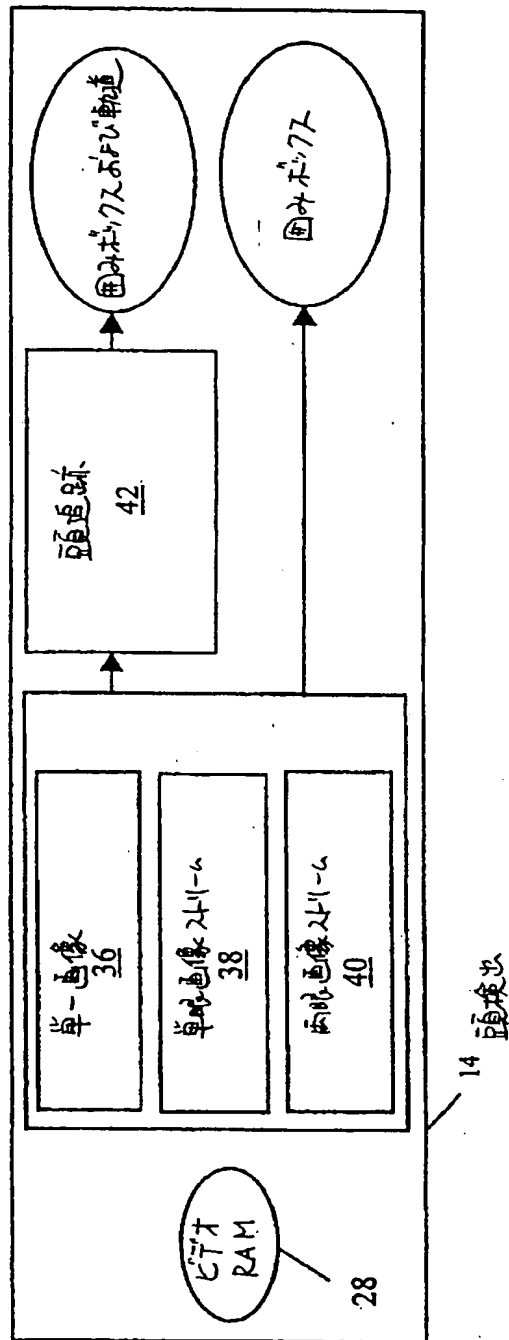
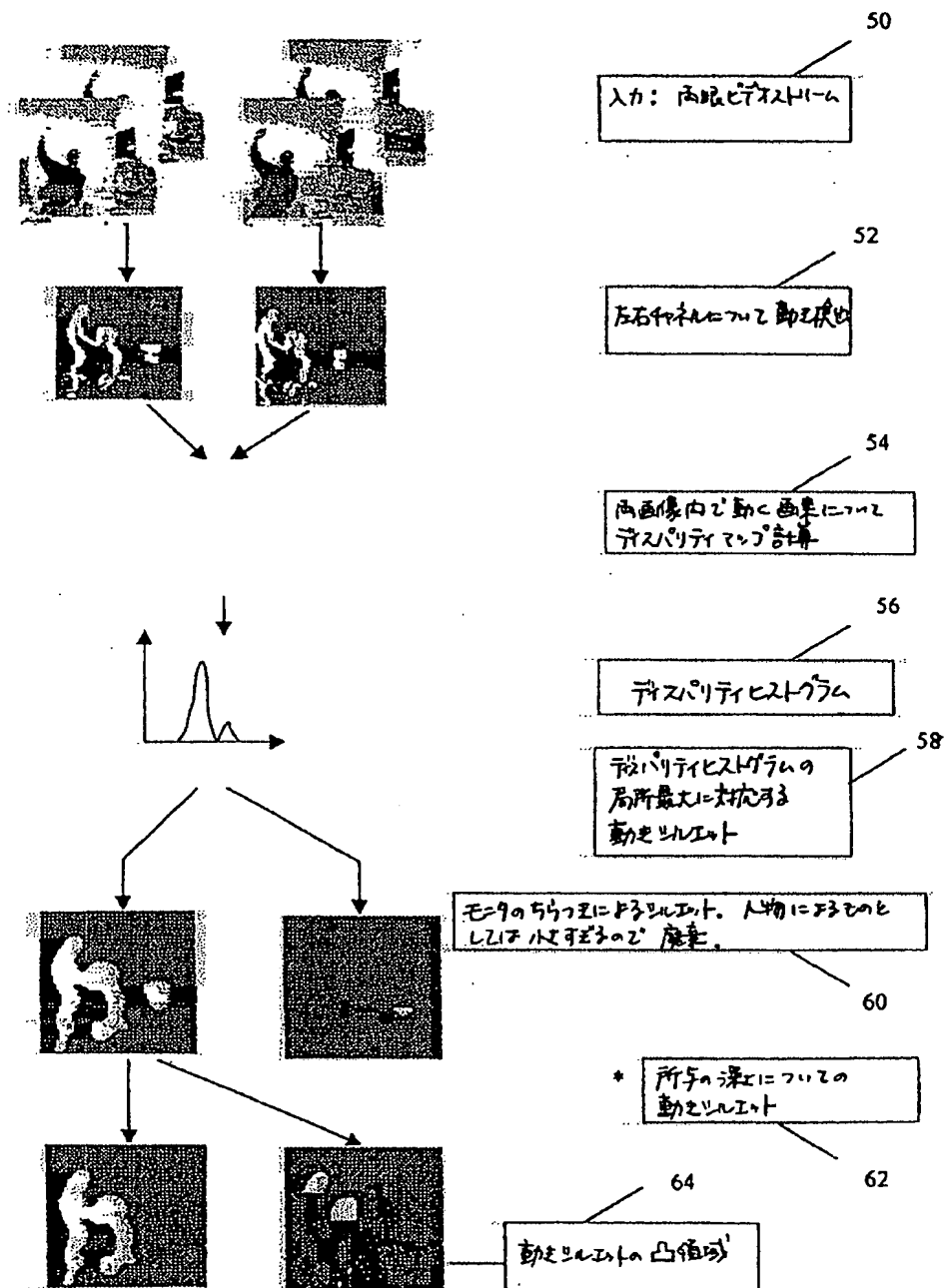


FIG. 3

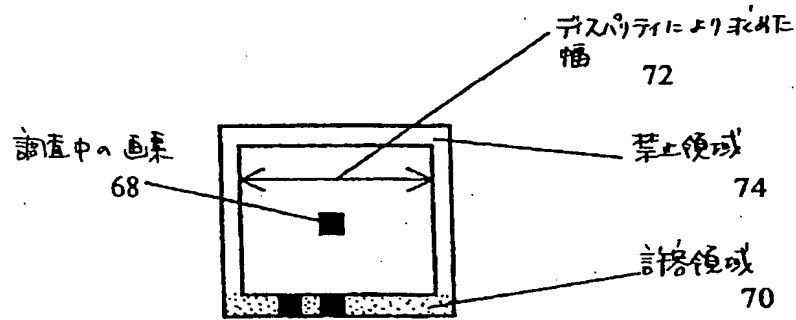
【図4】



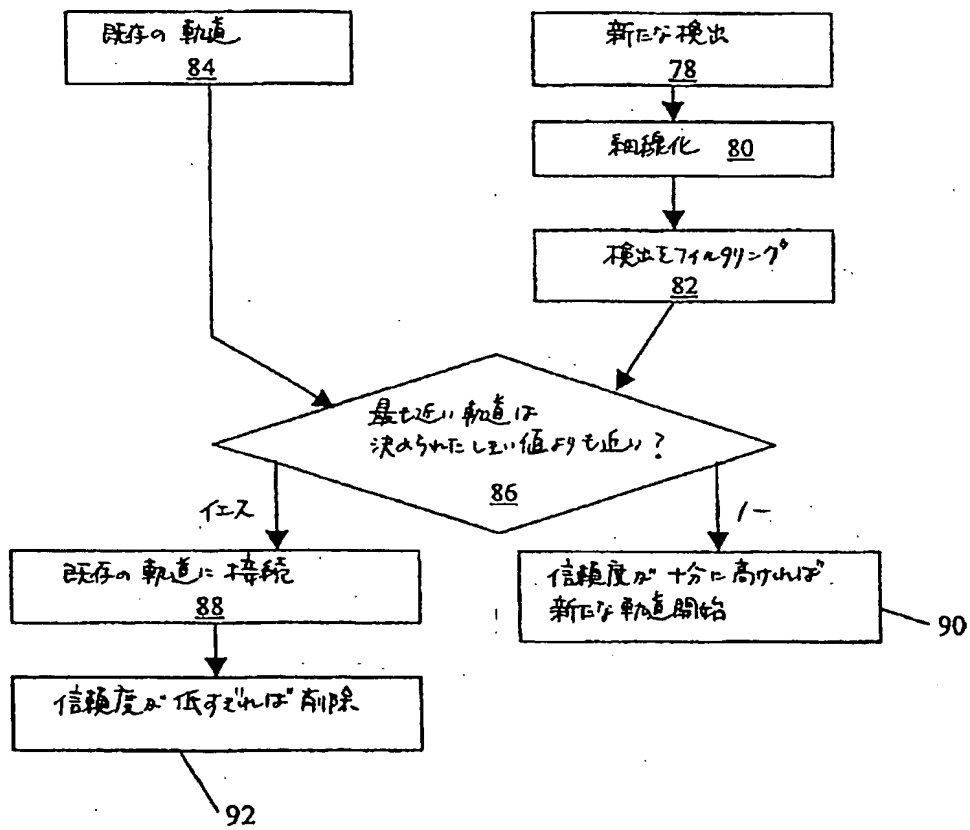
【図5】



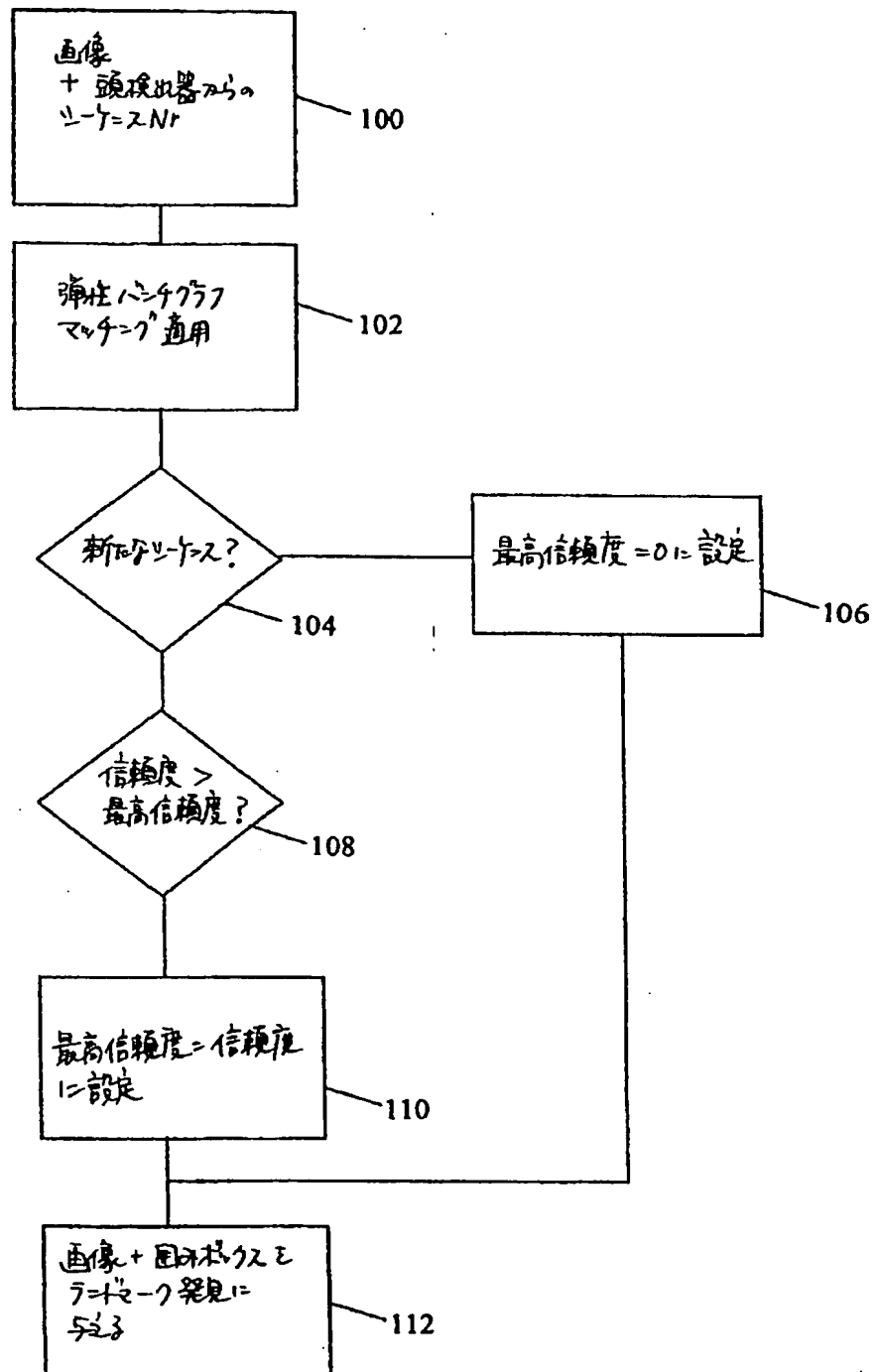
【図6】



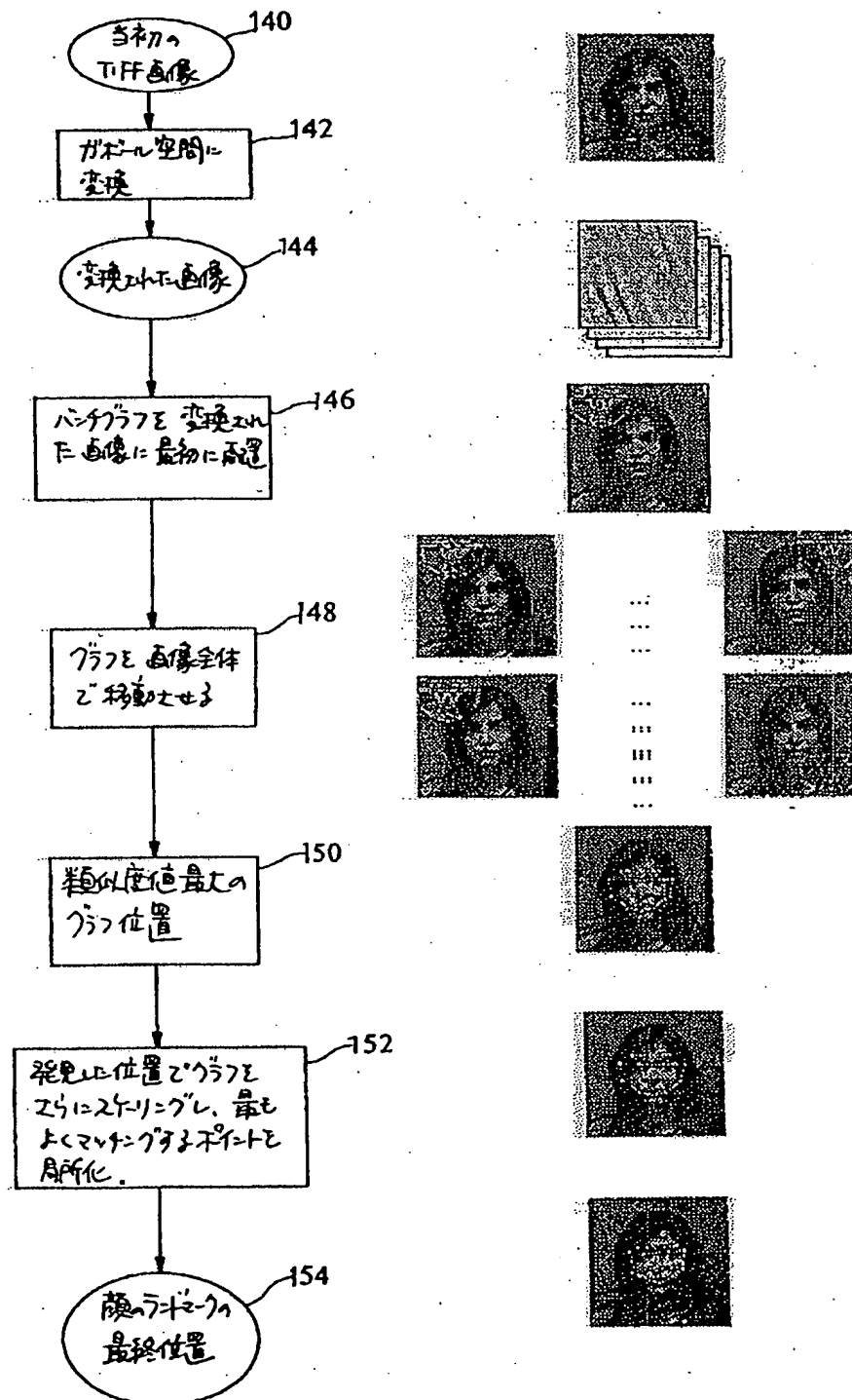
【図7】



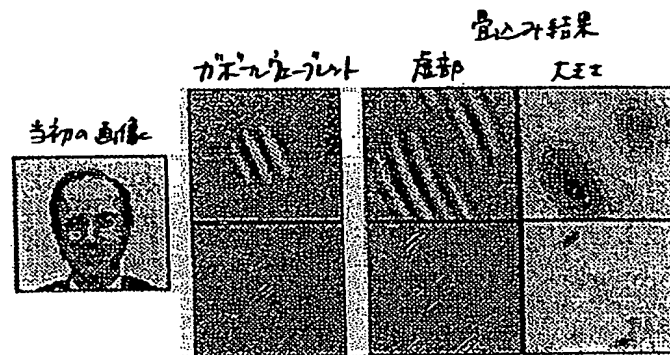
【図8】



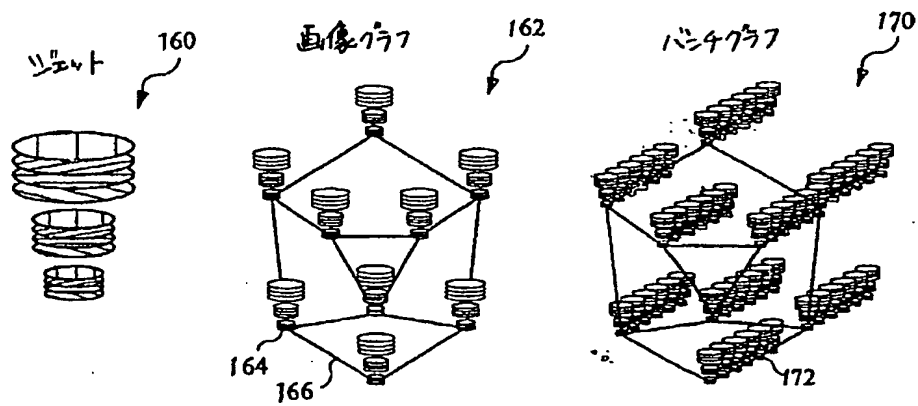
【図9】



【図10】



【図11】



【図12】

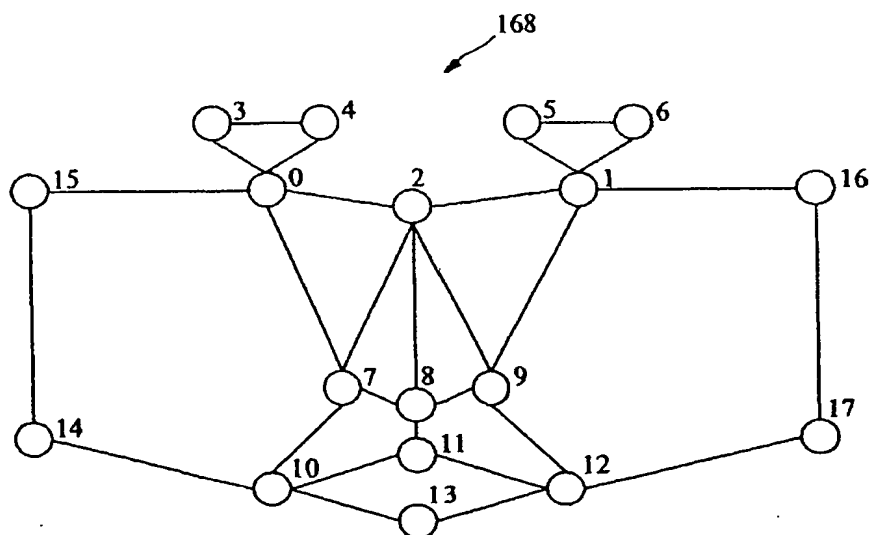
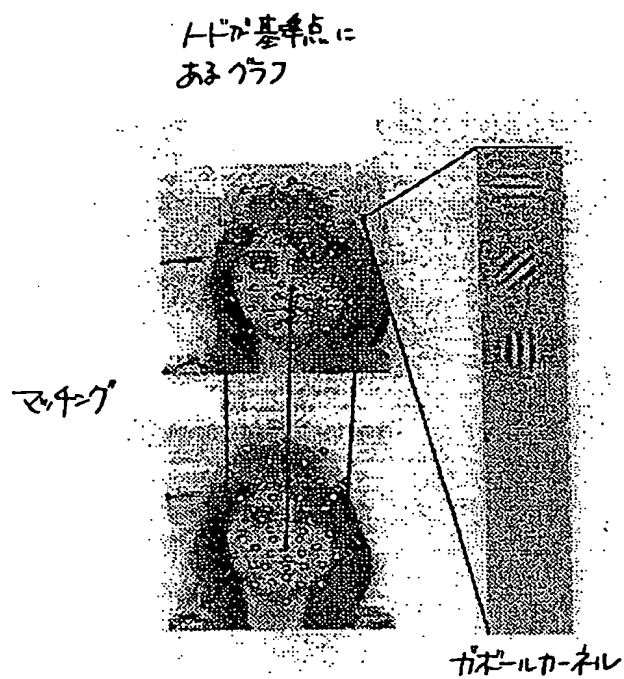


FIG. 12



【図13】



【図14】

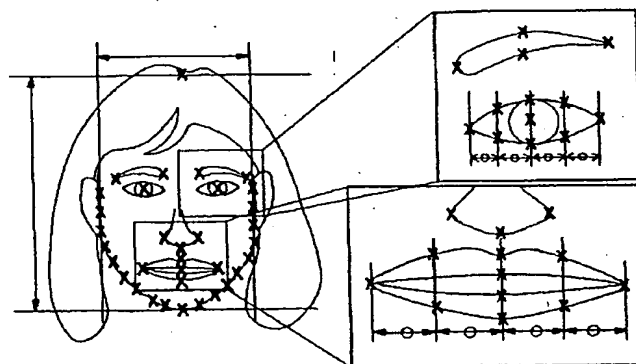


FIG. 14

【図15】

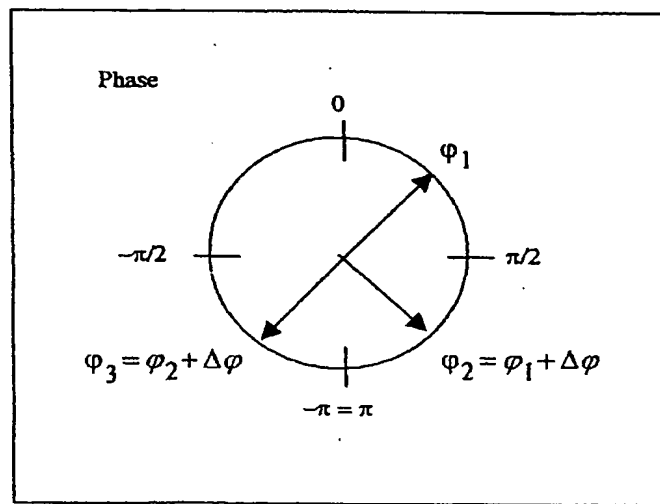
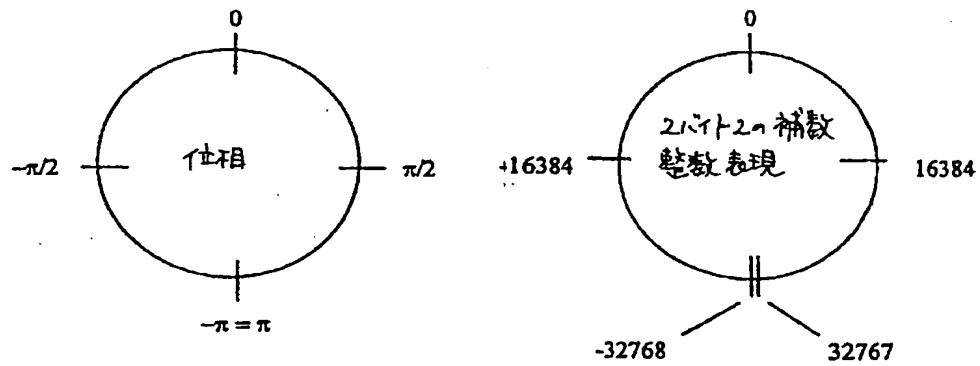
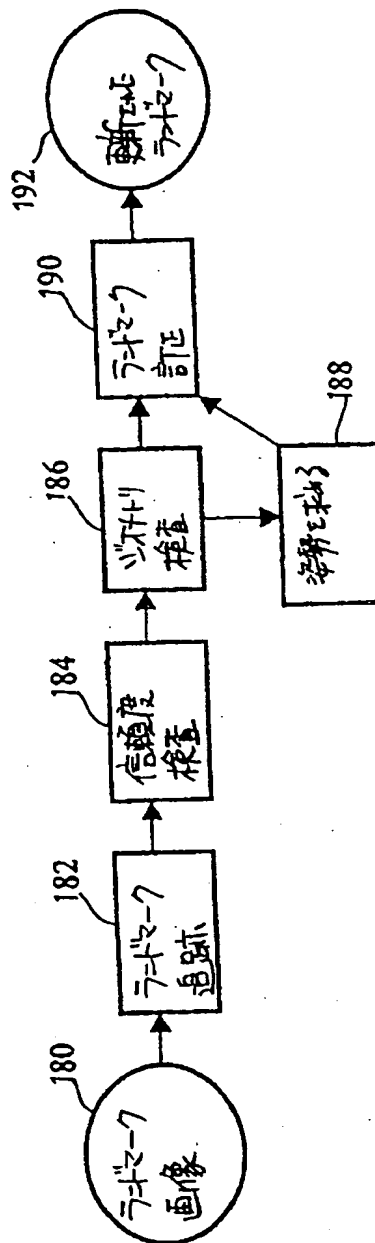


FIG. 15

【図16】



【図17】



【図18】



FIG. 18

【図19】

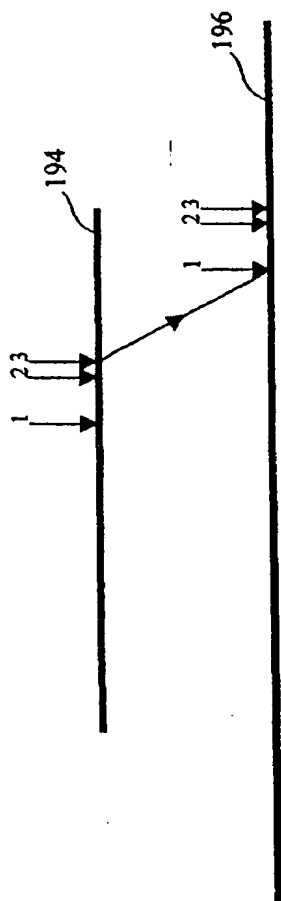
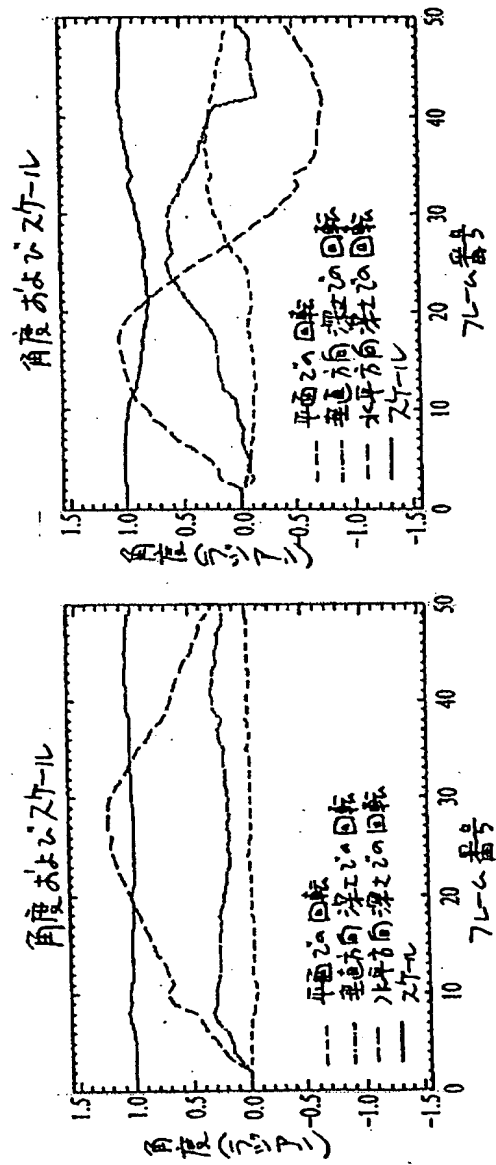
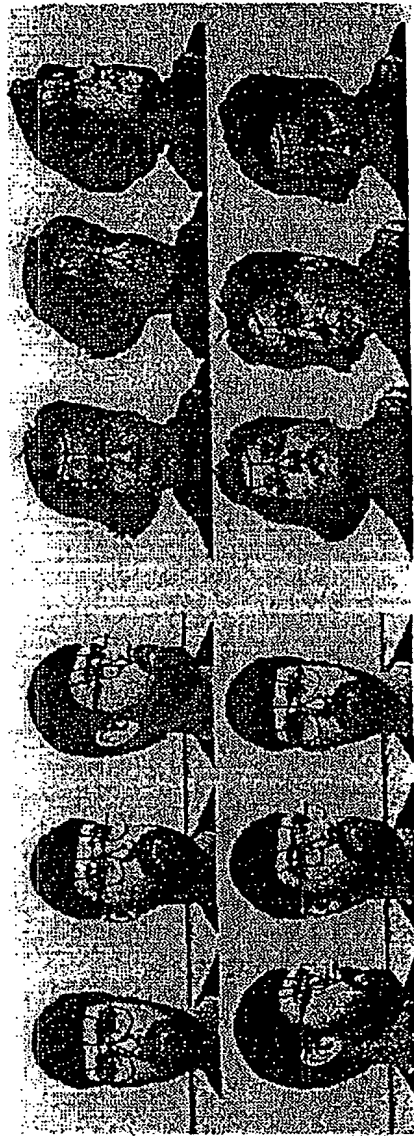
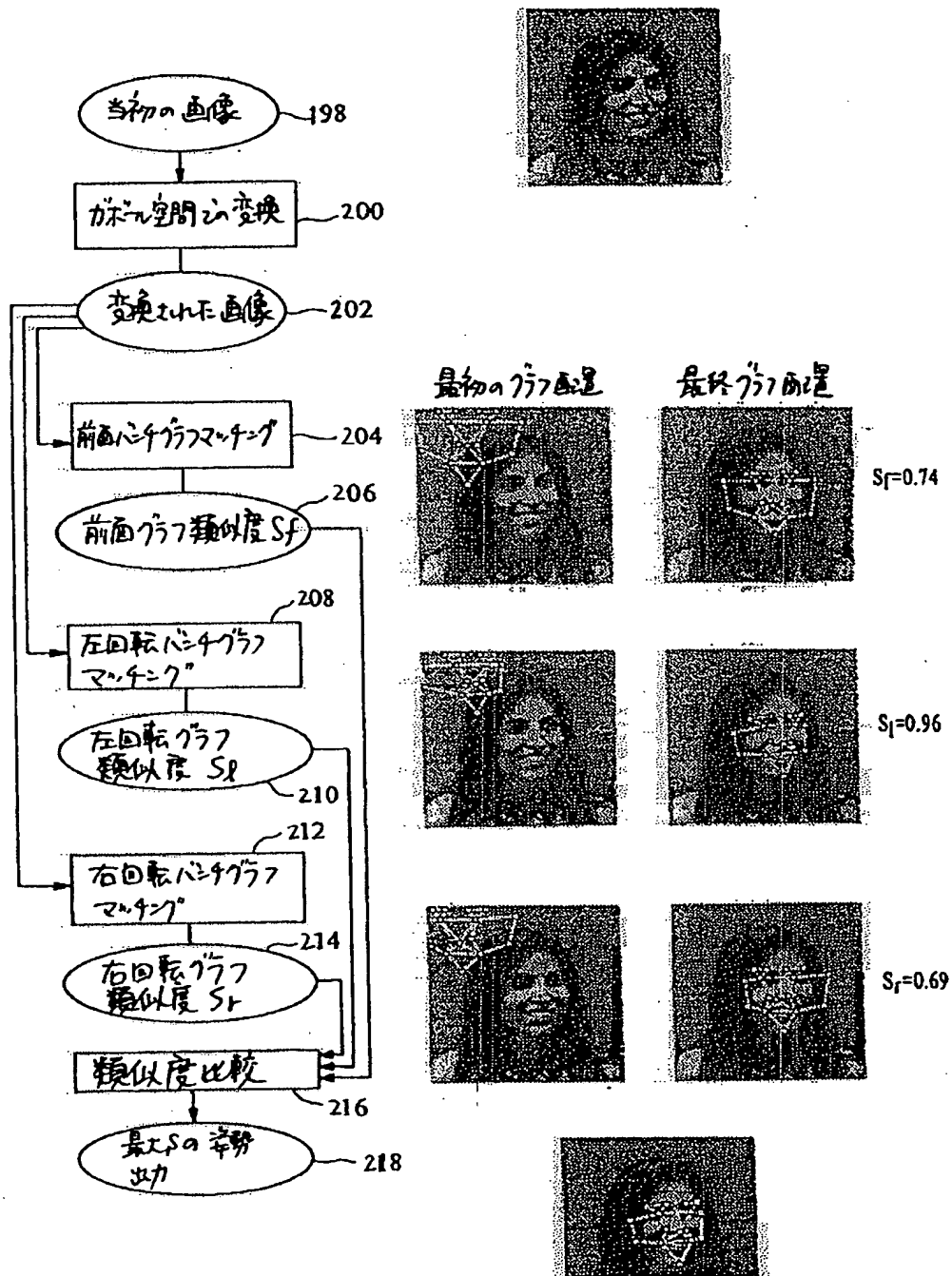


FIG. 19

【図20】



【図21】



【图22】

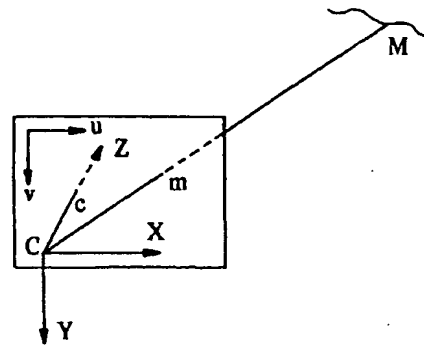
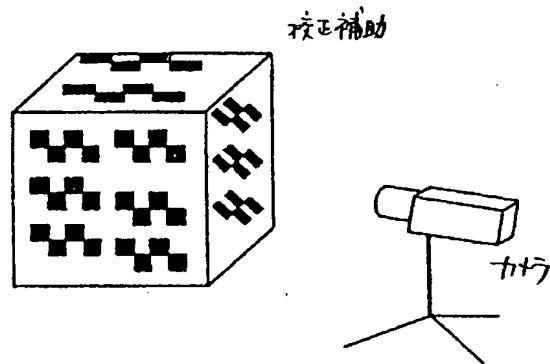
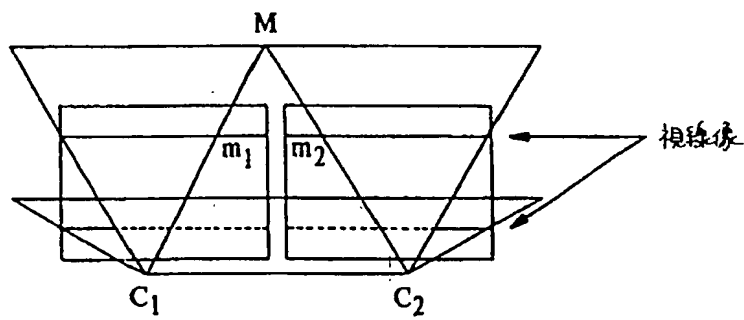


FIG. 22

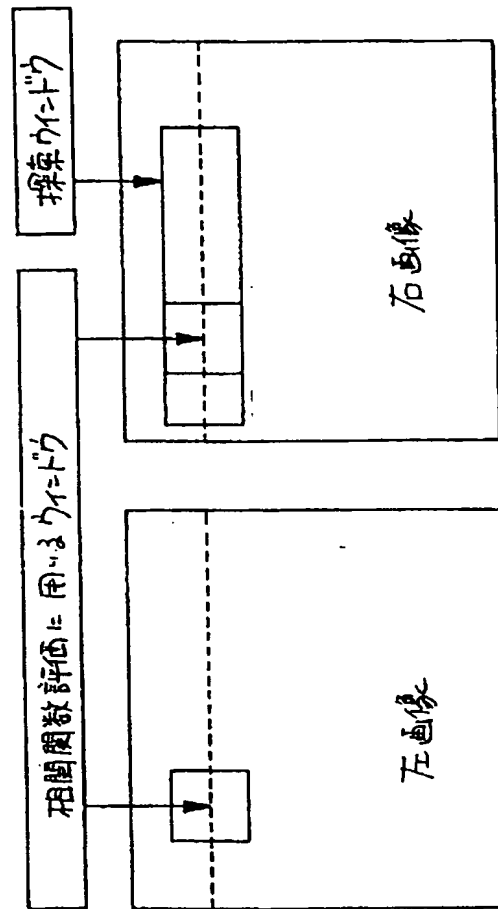
【图23】



【图24】



【図25】





【図26】

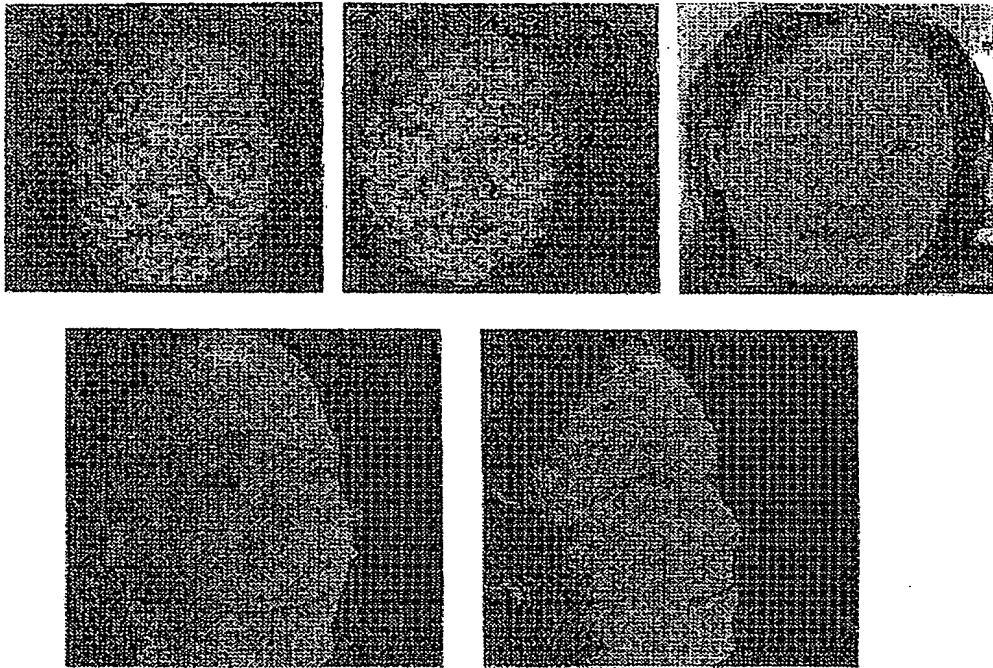
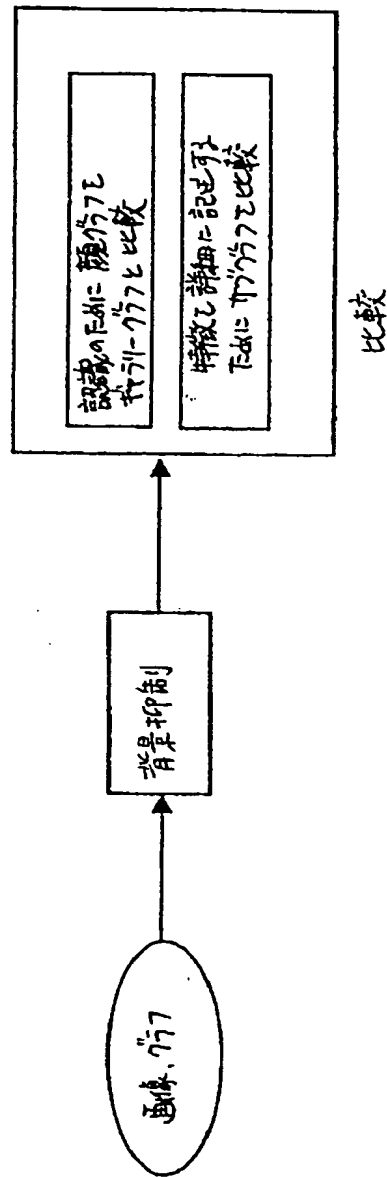


FIG. 26

【図27】



【図28】

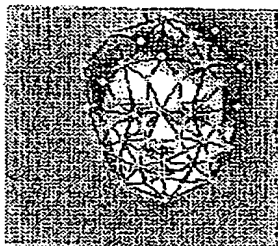


FIG. 28

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 6 G06K9/00		International Application No. PCT/US 99/07935
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 6 G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WUERTZ R P: "OBJECT RECOGNITION ROBUST UNDER TRANSLATIONS, DEFORMATIONS, AND CHANGES IN BACKGROUND" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 19, no. 7, 1 July 1997 (1997-07-01), pages 769-775, XP000698175 ISSN: 0162-8828 page 769, paragraph 2.1 -page 770, paragraph 3 page 772, paragraph 3.4 -page 773, paragraph 3.5 page 773, paragraph 4 -page 774 --- -/--	1,3,5,7, 8,11,23, 24
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		
<input checked="" type="checkbox"/> Patent family members are listed in annex.		
<b>* Special categories of cited documents:</b> "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document relating to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 17 September 1999		Date of mailing of the international search report 24/09/1999
Name and mailing address of the ISA European Patent Office, P.O. 5818 Patentlaan 2 NL - 2230 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer: Chateau, J-P

Form PCT/ISA/210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 99/07935

## C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WISKOTT L ET AL: "FACE RECOGNITION BY ELASTIC BUNCH GRAPH MATCHING" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 19, no. 7, 1 July 1997 (1997-07-01), pages 775-779, XP000698176 ISSN: 0162-8828	1
A	DE 44 06 020 C (ZENTRUM FÜR NEUROINFORMATIK GMBH) 29 June 1995 (1995-06-29) column 8, line 1 - line 40	1
A	WISKOTT L: "PHANTOM FACES FOR FACE ANALYSIS" PATTERN RECOGNITION, vol. 30, no. 6, 1 June 1997 (1997-06-01), pages 837-846, XP000690634 ISSN: 0031-3203 page 837 -page 839, paragraph 5	1

## INTERNATIONAL SEARCH REPORT

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
DE 4406020 C		29-06-1995	NONE	

## フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW

(72)発明者 エラジン, イゴー・バレリビッチ  
アメリカ合衆国、90007 カリフォルニア  
州、ロス・アンジェルス、セベランス・ス  
トリート、2636、ナンバー・エイ

(72)発明者 ノセラ, ルシアノ・バスクゥアレ・アゴス  
ティーノ  
アメリカ合衆国、90025 カリフォルニア  
州、ロス・アンジェルス、サウス・ウェス  
トゲート・アベニュー、1230、ユニット・エ  
フ

(72)発明者 ステファンズ, ヨハネス・バーナード  
アメリカ合衆国、90230 カリフォルニア  
州、カルバー・シティ、グリーン・バレー  
・サークル、6420、ナンバー・207

(72)発明者 ネベン, ハルトムット  
アメリカ合衆国、90405 カリフォルニア  
州、サンタ・モニカ、トゥエンティエイ  
ス・ストリート、2336、ナンバー・イー

Fターム(参考) 5B057 CA01 CA12 CA13 CA16 CG09  
DA11 DB02 DB03 DC36  
5L096 AA02 FA02 FA26 HA03 JA11